

DOCUMENT RESUME

ED 345 262

CS 213 329

AUTHOR van de Gein, Jannemieke
TITLE The Sense of Sentences: A Study into the Effects of Grammar Instruction upon Junior Writing = De Zin van de zin: Een studie naar de effecten van grammatica-onderwijs op teksten van jonge schaitchrijvers.
PUB DATE 17 May 91
NOTE 165p.; Ph.D. Thesis, University of Utrecht (Netherlands). Small print on some pages may not reproduce legibly.
PUB TYPE Dissertations/Theses - Doctoral Dissertations (041)
-- Reports - Research/Technical (143)
EDRS PRICE MF01/PC07 Plus Postage.
DESCRIPTORS Elementary Education; Foreign Countries; *Grammar; *Instructional Effectiveness; Sentence Combining; *Sentences; Sentence Structure; Theory Practice Relationship; Writing Research
IDENTIFIERS *Writing Development

ABSTRACT

This book presents the final report on a study into the effects of grammar instruction upon aspects of the writing of nine-year-old students. After an introductory chapter, the second chapter expounds the theoretical background to the study. The third chapter of the book introduces the experimental programs, the dependent variables, and the adjoining hypotheses. The fourth chapter, on composing, classifying, and constructing, deals with the courses competing in this study. The fifth chapter describes the instruments used to measure the effect of the treatment. The sixth chapter describes what happened, and how, and when, in the actual experiment; the seventh chapter presents tables of data. The eighth chapter discusses the results of the experiment: traditional grammar instruction did not enhance students' writing, but explicit instruction on grammatical peculiarities of written language can help eradicate specific troubles in writing. (Thirty-five tables of data and 3 student-drawn figures are included; 119 references and 2 appendixes of data are attached. There is a summary in Dutch.)
(RS)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED345262

THE SENSE OF SENTENCES

A study into the effects of grammar instruction
upon junior writing

Jannemieke van de Gein

CS 213 329

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

J. van de Gein

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☐ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

THE SENSE OF SENTENCES

A study into the effects of grammar instruction
upon junior writing

DE ZIN VAN DE ZIN

Een studie naar de effecten van grammatica-
onderwijs op teksten van jonge schrijvers
(met een samenvatting in het Nederlands)

Proefschrift ter verkrijging van de graad van doctor
aan de Rijksuniversiteit te Utrecht
op gezag van de Rector Magnificus,
Prof. Dr. J.A. van Ginkel
ingevolge het besluit van het College van Dekanen
in het openbaar te verdedigen
op vrijdag 17 mei 1991, des namiddags te 2.30 uur

door

Joanna Maria Petronella Helena van de Gein

geboren op 6 november 1954, te Vlaardingén

promotor: Prof. Dr. G.J. de Haan
Faculteit der Letteren, Rijksuniversiteit Groningen

co-promotor: Dr. G.C.W. Rijlaarsdam
Instituut voor de Lerarenopleiding, Universiteit van Amsterdam

Dit onderzoek werd mogelijk gemaakt met financiële steun van het Instituut voor Onderzoek van het Onderwijs (SVO) te 's Gravenhage.

ACKNOWLEDGEMENTS

The study into the effects of grammar instruction upon junior writing was a rather tremendous enterprise. Numerous creatures, great and small, assisted me and thus, be it literally or figuratively, alleviated the task I set myself.

My first thanks are due to the class teachers who allowed me to experiment with their students. They are, in alphabetical school order, Mrs. Nellie Schagen-van Latum of the Augustinusschool at Castricum, Mr. Bert van Lieshout and Mr. Kees Meuwese of De Beyenburgschool at Dordrecht, Mr. Ab Ploeger and Mr. Jos van Tilburg of De Bussel at Blokker, Mrs. José Heemstra of the Parkiet at The Hague, Mrs. Corrie Jongejan and Mrs. Josien Stassen of the Jules Verneschool at Hoorn and Mrs. An van Wijk of De Speel-Werkhoeve at Vlaardingen. I know I have been exceptionally lucky in meeting them and working with them. They were cooperative to the utmost. Their involvement and ready compliancy were no minor contributions to the success experimenting was.

Next, I want to thank my absolute top-of-the-bill experimenters: Ms Petra Plompen, Ms Caro Engel and Ms Renée Braams. Persevering happily and laughing doggedly, they were obliging and critical trainees and hard-working and loving teachers. In short, they are the best one could wish for.

Ms Ariane Blokzijl is the experimenter who came to my rescue when one of the magnificent three had to stop working. Willingly, she let herself be launched into this outrageous experimenter's life. I want to thank her for the excellent job she did.

An innumerable amount of data have been collected in the course of this study. Close to an innumerable amount of people helped to collect these. For doing this, I thank Karen Arens, Ineke Bakker, Marja van de Berg, Martie van den Bergh, Tatja Blok, Jan Boersma, Ingrid Bruining, Adri Bunte, Jack van de Burg, Lex Bruyn, Mette Caton, Hans van Deelen, Inès van Gassel, Annemarie

Gerbrandy, Carla van Ginneke, Annemieke van Gorp, Annemiek Hack, Annemarie van den Hoven, Marja van den Hurk, Mary van Koeverden, George de Koning, Monica Lamain, Ramon Mac-Donald, Yvonne Marchand, Mr. M. Mercks, Jacques Poell, Emilie van de Raa, Hester Ringnalda, Anneke Ruiter, André Sterk, Geke Uninge, Janneke Valbracht, Jeroen Vermeulen, Janet Weggemans, Erwin Witteveen, Ellemieke de Wolf, Cuneke Wouda and last but not least, Henka van Velzen and Rob Zethof of the Theresiaschool at Haarlem.

I am very grateful to Mr Gerrit Jan van der Duim for being such an outstanding commander-in-chief to these troops. He recruited and selected them, took great pains to train them and kept them under close surveillance. He actually spoilt me, for here they were, swarming the place in discipline and never, never bothering me. As if this wasn't enough, Gerrit Jan also took other cartloads of work out of my hands. Sometimes he simply insisted on doing tough jobs, sometimes I wanted him to do nice little jobs. I am not at all sure whether I realized just how nice and little these actually were. It seems as though, along with Mr. and Mrs. Boffin in Dickens's *Our mutual friend*, I always thought 'a Secretary to be a piece of furniture, mostly of mahogany, lined with green baize or leather, with a lot of little drawers in it'. Not so for Gerrit Jan. He has been my prop and stay and a true friend to boot.

It was a great relief having Ms Adri Bunte and Mr. Gert Rijlaarsdam next to me when designing and pilot testing the three experimental courses. Ms Janneke Valbracht and Ms Monique Prevaes assisted the both of them in criticizing the final versions of the courses I, and, in the case of the Direct Writing course, Gert and me, conceived. Their comment, though harsh at times, did a world of good to me and to the programmes.

Thanks to Mr. Edgar Tossijn and Ms Annemarie Gerbrandy working their heads off in transforming my unsightly scriblings, I was able to present neat protocols to the experimenters and lovely books to all students.

I thank Mr. Wim de Meier and Mr. Michael Vandeursen for all the printing jobs they performed so meticulously. I thank Mrs. Coby van der Ham and Ms. Irène van de Weetering, librarians both, for promptly providing the reading material I craved for. I thank all members of the secretariat of the research group for all jobs they assisted in. I want to single out Ms Linda Charité and Ms Nettie Koopman in particular. The former assisted me unforgettably enthusiastically at the very start of the project, the latter took professional care of the final manuscript.

Of my colleagues at the research group I want to single out Mr. Rob Schoonen and Mr. Henk Blok. They both lent me a patient ear and never seemed to tire of my seemingly endless need to talk matters over. And, what is more: they talked back!

Working with Dr. Mindert Eiting, another of my colleagues, has been a mixed blessing for both of us. Both of us very much wanted to turn the study into a success of sorts, but it turned out to be rather impossible to reconcile our very, very diverging views. Yet, I know I owe him quite a lot of what I know now

and so, in spite of the rather common fights we had, I want to thank him sincerely and very, indeed very much for all he did for me.

I seem to have noticed the special qualities of Mr. Amos van Gelderen, yet another of my colleagues, only at the close of the study. Heaven knows how he knew all the ins and outs of my study and heaven knows why he cared. But he did. His involvement and care meant a true refuge to me and made him very dear to me.

The same goes for Monique Prevaes and Renée Braams. By standing by me they encouraged me in no small way.

I thank the Institute for Educational Research in the Netherlands and the Kohnstamm Fund for their financial support. Also, I thank the members of my thesis committee, Prof. Dr. Henk Verkuyl, Prof. Dr. Wolfgang Herrlitz, Prof. Dr. Peter Jan Schellens and Dr. René Appel for their critical evaluation of this report.

I dearly thank Dr. Marjolein van Dort-Slijper for singling me out years ago. She meant me to articulate my thoughts on grammar instruction and did so very successfully. We thought and fought over innumerable issues, some minor, some fundamental. Challenging me critically and unremittingly, she helped me get a clearer picture of what it was I actually after. It is great working with her. Great fun, too.

Also, I sincerely thank Prof. Dr. Ger J. de Haan, who critically joined this project when it was well on its way. His support meant quite a lot to me. Yet, he disconcerted me severely by perpetually giving vent to his amazement over what I tried to do in this study and why. By thus provoking me to think over again and again what I thought, did and wanted, it seems he prominently stamped my present thoughts and future plans.

Most thanks, though, are due to Dr. Gert Rijlaarsdam, teacher, coach and sparring partner.

Jannemieke van de Gein
March 1991.

CONTENTS

ACKNOWLEDGEMENTS	V
1. INTRODUCTION	1
2. WHAT WRITERS NEED TO KNOW	7
2.1. Differences between oral and written language	7
2.2. What novice writers know	10
2.3. Unconsciousness at work: features of first writing	11
2.4. Developing a monitor for written language	15
2.5. Zooming in on sentences	19
2.6. Effects of grammar instruction upon writing	22
2.7. Conclusion	29
3. COPING WITH THE NUISANCE OF ACTUAL PRACTICE	33
3.1. What instruction ...	33
3.2. ... and what changes	34
3.3. Design of the study	36
3.4. Protocol adherence	38
3.5. Time on task	39
3.6. Students' progress	39
3.7. Regular curriculum	40
3.8. Summary	40
4. ON COMPOSING, CLASSIFYING AND CONSTRUCTING	43
The three experimental courses	
4.1. Pilot study	43
4.2. Direct Writing	44
4.3. School Grammar	47
4.4. Sentence Construction	52
4.5. The experimenters	59
4.6. Conclusion	60
5. FIVE DEPENDENT VARIABLES MEASURE FOR MEASURE	64
5.1. Pilot study	64
5.2. Writing fluency	67
5.3. Academic linguistic knowledge	71
5.4. Sentence awareness	72
5.5. Conciseness	75
5.6. Overall writing quality	76
5.7. Summary	77

6. ON THE ROAD TO NEVER-NEVER LAND: EXPERIMENTING	81
6.1. Method	81
6.2. Work in progress: quality matters	86
6.3. Quality of data (with Mindert Eiting)	91
6.4. Prelude to data analysis	94
6.5. Summary and conclusion	96
7. THE EXPERIMENTS: RESULTS	99
7.1. Writing fluency	101
7.2. Academic linguistic knowledge	104
7.3. Sentence awareness	106
7.4. Conciseness	113
7.5. Overall writing quality	115
7.6. Summary	116
8. DISCUSSION	119
8.1. Inviolable writing fluency	121
8.2. Kernels vs. universals: academic linguistic knowledge	124
8.3. The rise and fall of sentence awareness	127
8.4. Conciseness and overall writing quality	132
8.5. Aftermath	135
REFERENCES	137
APPENDICES	143
AUTHOR'S INDEX	149
SUBJECT'S INDEX	151
SAMENVATTING (Summary in Dutch)	153

INTRODUCTION

... a lower correlation between knowledge of grammar and composition than ... found between knowledge of history and arithmetic.

Boraas, cited in Bowden, The effects of formal, traditional grammar study on the writing ability of secondary school students.

This is the final report on the study into the 'effects of grammar instruction upon aspects of writing of primary school students' as was its official title. The study was carried out at the Research Group on Language Education ('Onderzoeksgroep Taalonderwijs') of the Centre for Educational Research ('Stichting Centrum voor Onderwijsonderzoek') of the University of Amsterdam and might be considered to be a continuation of the Wesdorp study into the goals and results of traditional grammar instruction. In his study Wesdorp found that language teachers and experts on language teaching thought it important for students in both primary and secondary education to have at their disposal terms and notions to discuss language and its usage. They proclaimed traditional school grammar an important part of the language arts curriculum as they thought this kind of grammar instruction did so provide students. And indeed, it did (Tordoir & Wesdorp 1979). Also, it turned out 91% of all responding teachers and experts found such grammar instruction to be important because of its facilitating foreign language learning; 82% of them thought the grammar instruction to be indispensable because of its salutary influencing language usage (Tordoir & Wesdorp *ibid.*; also cf. Wesdorp 1984).

Results of a first follow-up study suggested that the facilitating effect of traditional school grammar instruction upon foreign language learning in the first year of secondary education seemed unlikely as courses differed widely in set up and terminology (Tordoir & Damhuis 1982; also cf. Wesdorp *ibid.*). A second follow-up study, into the salutary effects of traditional school grammar instruction upon language usage of students, was scheduled but was never carried out. In preparing this study Wesdorp was confronted with serious theoretical problems. Or rather: he was not. Searching history for traces of

theory on the alleged beneficence of grammar instruction, he did not find a scrap of theory that could stand the light of day. Arguments and arguments only served to scaffold the stronghold of grammar instruction and, he granted, these did (and still do) so very efficiently. Wesdorp thought this to be a weird sort of conservatism, a remarkable display of resourcefulness, that, however, was rather useless as a basis for serious research (Wesdorp 1982).

But then these isolated, arbitrary 'theories' (if any) upon which hypotheses were based were a first moot point only. Second best were the hypotheses to be tested. Such a hectic lot of variables are involved in the acquisition of language skills; how on earth could any self-respecting researcher expect his pygmy treatment to be of any importance whatsoever? Moreover, measuring techniques were not by far sensitive enough to register the microscopic effects (if any) brought about by whatever treatment. For Wesdorp not just found theory, but measuring techniques to be very poor as well (Wesdorp 1981). Some of the instruments he had unearthed even made him think of the Dark Ages. He acknowledged that lack of an all-embracing theory on language skills impeded the making of adequate measuring instruments. After all, as each and every researcher had a theory of his own, each had measuring instruments of his own as well. No wonder these, more often than not, turned out to be dubious and defective and bore a medieval stamp.

Wesdorp (1983; also cited in Hillocks 1986) stretched his point even farther by proclaiming that all writing researchers carried on almost blindly. Those who were after the beneficent effects of grammar instruction were a case in point only. He claimed that all writing researchers ignored previous studies, refused to search carefully and systematically through earlier work, and were apparently not even prepared to start building anything like a theory. This is an admittedly harsh condemnation of all writing research, this one included.

Nevertheless, Wesdorp still felt confident about research. He recommended the smallest small scale experimental research with very explicit though be it ever so modest expectations. Also, he thought that each measuring instrument, however questionably and defectively exploring whatever detail of language skill, still meant one step in the right direction. And so, fragmented, fashionable and detailed though they may be, there are studies that show some sort of relationship, or, as Hillocks put it: 'Lack of consistent research, however, does not imply that experimental research has been random' (Hillocks *ibid*: 96). Pecking through fashionable ideas, limited instruments and delicate variables of their colleagues, plodding through very diverse results indeed, prejudiced researchers are working their way towards some sort of theory. Thus, around some of the issues in writing research a growing body of knowledge is gathering, albeit tauntingly slowly. One of the issues is the relationship between traditional school grammar and writing. One of the researches is this one. Right from the start, then, the study into the effects of grammar instruction upon writing of primary school students was known to be a perilous enterprise. And indeed, it was.

This study contrives to show that familiarizing students with the terms and notions traditionally taught in school grammar does not enhance students' writing. This is not because it is grammar but because it is the wrong kind of grammar. As this kind of instruction does not teach students what they need to know in order for their writing to improve, whatever knowledge there is, does not transfer to writing. For, as this study has it, junior writers definitely have to be initiated in some of the formal characteristics of written language. After all, and as a matter of course, the peculiarities of written language are not covered by the tacit grammatical knowledge of language users. Instruction in the peculiarities of written language should, very emphatically, not cover what are or are not syntactically well-formed structures in language, as even apprentice writers, however new to this job, know to a tee what are and are not. In this respect, then, there is nothing essentially the matter with their writing. Young writers might, of course, their handwriting still being slow, in haste leave out words that should have been there. They might, being the reluctant rewriters they are, never notice the odd deviant structure. Attention may falter occasionally, resulting in blended structures. They even might experiment with this newly acquired language and thus produce awkward structures in stilted wording. As, however, the faculty to use language fluently resides in the tacit grammatical knowledge of language users, writing fluency is fine and need not be bothered with¹.

Central in focus is the sentence, an outstanding curiosity of written language and for writers new to the trade a stumbling block instead of a building block. Sentences, that is to say: the fully-fledged sentences of written language, do not come naturally to language users, whereas at the same time they are an essential scaffold to the writer, structuring content on its lowest level. Hence, if one of the characteristic aspects of writing is wittingly giving form to meaning, the sentence is a target unit in planning, formulating and rephrasing one's thinking.

As sentences are writers' stepping stones, apprentices need to get acquainted with their basic building blocks and with the slots a basic sentence vouchsafes writers to expand and elaborate meaning. This is where terms and notions traditionally associated with grammar instruction in school comes in.

This study, then, turns to grammar instruction only as a means to enlighten students where they need be enlightened. It does not, as has been done previously, and that to no avail, ply students with all paraphernalia that traditionally comes in the suite of school grammar. There seems to be no need to do so at all. That is, not as far as junior writing goes.

In this respect this study is decidedly alike the twofold study Dirksen and Schuurs conducted (Dirksen 1990; Schuurs 1990). They carefully analyzed essays of twelve-year-olds and found there were errors and errors. Some, they maintain, are accidental, some are not. The accidents, no errors in the real sense of the word, are caused by momentary attention dips. Such accidents are, e.g.,

1. As if it could be bothered with!

words that appear twice over or not at all, awkward structures, unwittingly funny, and agreement errors (see especially Dirksen *ibid.*). There is no need to worry overmuch about these errors: have a writer reread his prose carefully and his tacit grammatical knowledge will take care of them.

Dirksen and Schuurs, quite rightly, profess the tacit grammatical knowledge of writers does not, and cannot, take care of other errors. Their studies have what they termed 'interaction errors' for a pivot. These concern erratically gapped sentences and erroneous conjunctions and references. Such errors, Dirksen and Schuurs profess, find their origin in the writer's lack of distance: writers may forget their prose is meant for a reader who needs to be informed meticulously or else the message won't get through. They hoped grammar instruction would eradicate these errors. But it did not, or so it seems. The present study also concerns errors tacit grammatical knowledge cannot cope with. Here they are errors that find their origin in a lack of expert knowledge on the fully-fledged sentences in writing. The study hypothesizes grammar instruction will result in such expert knowledge. And it does, or so it seems.

Though all evidence as to its beneficial effect are still lacking, instruction in traditional grammar is, at least in The Netherlands, still widely held to be a fine means to enhance students' writing skill and thus this kind of instruction, cloaked conventionally or in modern out fit, usurps a substantial part of the regular language arts curriculum (cf. Bos & Oostdam 1985; Zwarts 1990). Contrarily, declared opponents of this end of grammar instruction hold that in order to ameliorate their writing, write is what students should do.

This study puts both creeds to the test and confronts them with a third creed to be tested, viz. that student writers need to be provided with rules on the peculiarities of written language. Though these rules pertain to writing only, they will be presented to students in a no-writing (or: a-rhetorical) setting, that is to say: in a setting equal to that of traditional grammar instruction. The study maintains that as the rules concern written language, they will transfer to writing and as they concern, more in particular, the sentence, they will effect writing of a superior quality. Neither a course in traditional school grammar nor a course in writing will thus affect writing as neither provides students with the knowledge they so badly need.

Proclaiming sentences pivot in grammar instruction and writing does of course not imply that they are the be-all and the end-all in writing instruction. There is more junior writers need to learn about. But that is not what this study is about.

Neither should it be taken to mean that prose should be frozen into full clauses and in full clauses only. In order to effectively manipulate the form of meaning, though, writers need to know what the options are. When reading a piece of prose such as (1.1)

- (1.1) I had a friend called
 michiel but alas
 he moved house but
 we regularly write
 letters and this way he stays
 a friend all the same.
 (6/21/3; Silvester)²

Ik had een vriend die
 heete michiel maar jamer
 genoeg is hij verhuist maar
 wij schrijven regel matig
 brieven en zo blijft het toch
 nog een vriend.

one might be moved and wont to credit the writer with an exquisitely budding feeling for poetry. One should however do so if and only if the writer determinedly chose to write poetry. This writer didn't. This is simply all he is up to. It may be effective here but it is very ineffective almost anywhere else, and he needs to know this. Once armed with a repertoire, able to choose, he is free to do as he wishes.

This report comes in seven chapters, this one excluded. Chapter 2 expounds the theoretical background to the study. First it highlights major differences between oral and written language: speech is a naturally acquired way of communicating with public present and participating; writing is a highly formalized, lonely, way of communicating with an absent public. Next, sections 2 and 3 deal with the knowledge the most tender aged of all writers possess (a lot) and use (a lot less). Though it knows writing differs from speech, the tiny essays the small fry write do not bear evidence to this at all.

Section 4 focuses on a particularly vital change in writing development: attention shifts to the language writing comes in. This attention shift causes malaise. Novices, no different from any other language user, are not possessed of expert knowledge on written language: standard writing does not come naturally. Section 5, emphasizing writing may well be said to be a foreign language, zooms in on one of the most outstanding features of standard written language: the sentence. As both sections 4 and 5 bring out that junior writers want expert knowledge on this foreign language, section 6 quests for the best type of instruction to provide them with this expert knowledge and comes up with a 'sentence construction'-type programme. This, argues the concluding section of chapter 2, will help students structure content in writing.

Chapter 3 introduces the experiment. First, the three experimental programmes, five dependent variables and the adjoining hypotheses pass in review. Next, the design of the experiment, a randomized block design, is dealt with. The remainder of the chapter relates what measures were taken to ward off the evil lying in wait to slaughter this vulnerable creature. Protocol adherence of experimenters, students' progress as well as their time on task and, finally, the regular curricular language arts programme were all controlled for.

2. Procedures in translating student texts were as follows: the Dutch version renders the exact student text, erroneous spelling and punctuation included, starting on a new line whenever the student did so. In the English translation I have tried to stay as closely as possible to the original text except for the spelling errors. As I feared 'translating' errors might very well turn the English version into something very weird, not at all resembling anything English speaking nine-year-olds would write themselves, I chose to leave out all spelling errors in the English version. Footnotes will alert the English reader to passages that proved to be hard to turn into English without harming the noteworthy original.

Chapter 4, on composing, classifying and constructing, deals with the courses competing in this study. One, obviously, is a sentence construction course. Its focus is the patterns underlying sentences and the pattern slots that cry for detail.

The effects of this course are contrasted to a course in traditional school grammar, dealing with nouns, verbs, subjects, plural and the like, and one in writing. The writing course touches on paragraphs, dialogues and heads, bodies and tails of texts. It has students write and prewrite and rewrite. It has them discuss texts as well.

All courses have been pilot tested. Section 1 relates what exactly this pilot study entailed. Sections 2, 3 and 4 present the ins and outs of courses. Section 5 presents the Misses engaged to teach all courses. Section 6, the final section of the chapter, neatly summarizes the essential characteristics of the three experimental treatments. Hypotheses end the chapter: what course will yield what result?

Writing fluency will not (and indeed: cannot) be affected by treatment. Academic linguistic knowledge (such as recognizing parts of speech), sentence awareness, conciseness and overall writing quality will be, though. Chapter 5 concerns the instruments used to measure effect of treatment. It provides examples for test items. To illustrate the workings of some of the measures, these are set at two essays. Tables summarize essential information on the test battery.

What happened, and how, and when, in the actual experiment is to be read in chapter 6. The chapter fraternally presents mundane details of the doings of Misses and statistical analyses of the doings of their students. These are preceded by matters pertaining to method. Psychometric information on tests and measures follows suit. The chapter finishes by presenting the way data analysis will be performed.

Chapter 7 is a cascade of tables: it presents results. Four figures illustrate the experimental effects. These bring out whether hypotheses held or not. Some did. Some did not.

The final chapter of the report, chapter 8, engages in a discussion of results. It argues how come the students that studied traditional school grammar do not know what verbs are, it discloses how the sentence construction programme helped students get to grips with exactly what kind of unnatural sentences and it tries to unearth what might spur apprentice writers to elaborate concisely: the writing programme or the sentence programme. The final section of chapter 8 views over the aftermath.

WHAT WRITERS NEED TO KNOW

'Let us see the cause of this cursed clamour,' said Front-de-Boeuf; 'here is a letter, and, if I mistake not, it is in Saxon.' He looked at it, turning it round and round as if he had had really some hopes of coming at the meaning by inverting the position of the paper, and then handed it to De Bracy. 'It may be magic spells for aught I know,' said De Bracy, who possessed his full proportion of the ignorance which characterised the chivalry of the period. 'Our chaplain attempted to teach me to write,' he said, 'but all my letters were formed like spear-heads and sword-blades, and so the old shaveling gave up the task.'

Sir Walter Scott, Ivanhoe.

This chapter expounds the theoretical background to the study into the effects of grammar instruction upon aspects of writing of primary school students. Section 1 concerns the characteristic differences between oral and written language. Section 2 is concerned with the grammatical and textual knowledge students possess at the onset of academic language instruction. The third section discusses content, organization and syntactic aspects of text of the youngest of young writers. The fourth section does the same for writers old or mature enough to exert conscious control over their writing. The latter section highlights the attention shift of these writers while monitoring their writing. It purports that junior writers thus obviously attempting to differentiate writing from speech determinedly try out hypotheses on written language and argues that these desperately seeking students should be educated so that they no longer have to guess wildly. Section 5 zooms in on a point in case: sentences. Section 6 presents and discusses research into the effects of grammar instruction upon writing. Section 7 concludes the whole of the chapter.

2.1 Differences between oral and written language

Language comes in many guises. To name two: there is language for the self and language for others. Language for others comes in oral and written texts. Both come in marked and unmarked formats. This report exclusively deals with language written for others. This section narrows matters down even further as differences between speech and writing will be discussed by focusing on the unmarked formats they come in only. This is a gross oversimplification of matters. After all, there is no unidimensional distinction between written and oral texts. A letter to a close friend is a typically rambling conversational piece of prose whereas a public talk is nothing like a conversation. Not in its

unmarked guise, that is. However, these finer multidimensional distinctions between writing and speech will not be considered at all.

In its unmarked guise, writing is an almost exact opposite of unmarked speech. The latter is generally taken to be an informal, conversation-like variety of language characterized by its familiar style. In its unmarked format speech comes naturally to language users. It is the first type of language a child becomes familiar with. Writing on the other hand is usually associated with rather edited, formal and polite varieties of language that seem to come far less naturally.

There is a rather substantial body of research on the differences between writing and speech in their unmarked shapes. Purely linguistic differences, concerning the structure of individual sentences, as well as differences on the level of the text as a whole have been studied. These studies have brought out that written language, generally speaking, tends to be syntactically more complex, has a higher density of new information and is better organized whereas speech is characterized by incomplete grammatical constructions, interruptions, hesitations and afterthoughts. Recorded speech is something decidedly different from writing and especially renowned for its 'general vagueness, resulting from the loss of interaction and gesture that would have accompanied the original utterance' (Herriman 1986: 166).

These differences are due mainly to the very characteristics distinguishing writing from speech. First of all, speech is volatile and thus limited by such mundane matters as memory span, attention and concentration. Writing lasts, so it enables the writer to deliberately attend to form and content, inviting preciseness and explicitness hardly ever equalled in conversational speech. But then, of course, speech need not be equal to writing in this respect at all. Partners in conversation always share a body of knowledge, so a lot of what should be mentioned to strangers may remain unmentioned or implicit. Besides, partners in conversation are informed through non-linguistic means as well: hands and faces speak volumes. Finally, partners are able to interrupt each other when they do need extra information. They do not even have to do so by word: raised eyebrows, eyes on stalks, gesticulation are enough. By simply being there, all partners in conversation are actively involved in what is being said. They jointly participate in constructing 'text', all cooperating to make it an organized whole. Even when chiefly listening, they may be heard to contribute by, e.g., interrupting the speaker or adding to what he just said. These interactive dynamics of speech are enlivened all the more by speakers themselves, as they use intonation, repeat words and phrases, and clarify themselves by going on after short pauses: 'So...', 'And...', 'After all...'; they ask for confirmation: '...surely?', '...don't you think?' and signal they are rounding up: 'Well, ...', 'All in all, ...'. In short, casual speech is a dynamic stream of words, a text built on the spot, to which information is added linearly without deliberate planning ahead. This does not mean language users do not in any way plan or evaluate, in other words: monitor, their conversational speech. They do. Speech errors such as spoonerisms, i.e. tips of the slung such as 'whalt misky', in Dutch unequalled

in 'De schand in het Breveningse Hoerkous', or blending errors, i.e. merging two different syntactic structures into one (e.g. Eng: He behaved as like a fool; Du: Zij draagt een geel jurkje aan) are generally taken to evidence the forward monitoring activities of speakers: they bear witness to speakers planning the syntactic structures of and the words in their utterances. Moreover, spontaneous self-repairs, i.e. speakers correcting blending errors or spoonerisms spontaneously indicate monitoring in retrospect, or evaluation. Monitoring such as this as a rule seems to be associated with the tacit knowledge language users acquire naturally, almost accidentally, and without any instruction.

In all probability language users do not monitor their utterances for grammatical correctness only. Other monitoring activities are, e.g., those for intended meaning and for appropriateness.

In writing there is no partner to interact with. Writers typically try to communicate with an absent and only vaguely familiar public, single-handedly constructing a meaningful text that has to be completely clear without all the paraphernalia that accompany speech. Writers have to make do with as scanty instruments as sentence constructions and punctuation marks. This does not mean writing is a hopeless venture. Writing, after all, can, and, as a matter of fact: should be constructed far more deliberately than speech is, word by word, sentence by sentence and paragraph by paragraph. Guided by their notions about the public they try to make sense to, writers turn into readers, pondering over the text at length, leisurely contemplating and reorganizing information, moulding the text, planning sentences and paragraphs. Writers, though, should not just read their own texts; they should read these with the eyes and mind of the strangers it is meant for, blank as far as the text at hand is concerned. Writers are their own partners, involved in as well as distanced from their own text, not merely producing language but rather consciously controlling it.

Summarizing, this section sketched a rather caricaturally but nonetheless true picture of the essential differences between speech and writing. Casual speech is a naturally acquired way of communication, generally not demanding deliberate planning activities as oral text is created on the spur of the moment in a shared effort to communicate. It resembles writing as one and the same linguistic and communicative competence is involved. It is opposed to writing as the latter is a highly formalized way of communicating, requiring deliberate planning and evaluation activities, written text growing out of the writer's lonely struggle to make sense to whomever it may concern.

1. Recorded in its most complete version in Battus 1981: 93.

2.2 What novice writers know

Counting out nursery school, children generally start their academic career at the age of six. At that age they are in full command of the grammar of their mother tongue. Evidence abounds: six-year-olds are able speakers and hearers, able quarrellers, pleaders, discussers, soothers, commanders and jokers. They know how to boss their baby brother around in 'motherese' and some young children are downright chatterboxes. Moreover, they are able to judge the well-formedness of sentences and when asked to correct a grammatically deviant sentence they will, as a rule, do so (see, e.g. Sinclair et al. 1978; Hakes 1980; Tunmer & Bowey 1984; Tunmer & Herriman 1984; Felix 1987; Galambos & Goldin-Meadow 1990). However, in detecting and correcting structurally offensive sentences, they have been shown to go by their tacit grammatical knowledge only. To wit: they are unable to explain what is wrong with the deviant sentences they so easily recognized and corrected (Galambos & Goldin-Meadow *ibid.*) and they are unable to repeat a syntactically deviant sentence: they correct it unawares, probably not even conscious of its syntactic offensiveness (Bowey 1986a, 1986b). Language users consciously controlling language, e.g. when treating the language system as an object of thought, do so through mediation of cognitive systems other than the language system ('problem-solving cognitive structures', Felix *ibid.*: 155). Six- to eight-year-olds are widely held to be far too immature cognitively to exert any such conscious control over language.

Instruction in writing and reading for real starts at school. However, children as young as three or four years old have been shown to be familiar with written language, albeit the texts they write do not at all resemble anything conventionally recognized as writing. In texts of such 'early writers' there may be words, made of rudimentary letters, some right side up, some upside down. Besides, there are all kinds of dots, coils and twists to behold. Drawings are an integral part of such texts. In order to 'read' these texts the writer's help is needed definitely. Also, one might have to turn around the page a number of times: text is spread over the page topsy-turvy, from left to right and back again (see, e.g. Dyson 1983; Sulzby 1986).

Well-read to four- and five-year olds are claimed to possess and use knowledge on genre-specific syntactic peculiarities of written language, especially narrative. They are aware of its being syntactically more integrated, meaning they do not add information linearly but embed it when recreating written narrative, 'reading' a narrative they have just been read themselves to a doll from a picture book. Also, they know what words typically belong in the context of a written narrative. Finally, they are able to involve their public in the narrative while at the same time maintaining its essentially decontextualized nature (Purcell-Gates 1988). To preschoolers living in a literate world this knowledge comes naturally, that is to say: they are not taught any one of these things explicitly. So this knowledge is as tacit as is the grammatical knowledge of these children.

2.3 Unconsciousness at work: features of first writing

Content and organization

When at school children first of all have to learn to write. They have to learn the alphabetic code and they have to learn to draw letters. Letters ascend and meander, there are curves and dots to be made. It takes a lot of time for handwriting to get fluent at all and even when children are relatively skilled, their writing is agonizingly slow and always recognizable as that of a child. Conventional writing only starts if this skill is more or less automatized, i.e. when students are about seven years old (see, e.g. Britton 1970; Burgess et al. 1973; Rosen & Rosen 1973; Graves 1975). Automatized here does not imply mastery. It means the activity does not usurp over-much conscious attention.

Texts of novice writers are a far cry from conventional writing, still. Writers new to the trade produce very little text. Of course theirs is a small world and their lack of anything like encyclopedic knowledge is a serious barrier to their producing text of substantial length. But then, texts written on subjects not asking for any such knowledge are extremely short as well. Two examples will go to illustrate this. Text (2.1) is written by Walter. He writes on his favourite pet. Rob wrote (2.2), a description of one of the mugs in figure 2.1.

(2.1) it is a sheepdog
it has a green fleacollar
it is a frisky dog
(5579/6; Walter)

het is een herdershond
het hebt een groen vooiebantje
het is een spelse hond

Figure 2.1 Drawing of mugs. Testees had to choose one and describe it.



2. For procedures followed in translating students' texts, see chapter 1, note 2.

- (2.2) There is an elephant on it
and there are little stars
it is small and it has a handle
(4/13/29/M3; Rob)

Er staat een olifant op
en er staan sterretjes
hij is klein en hij hebt een handvat

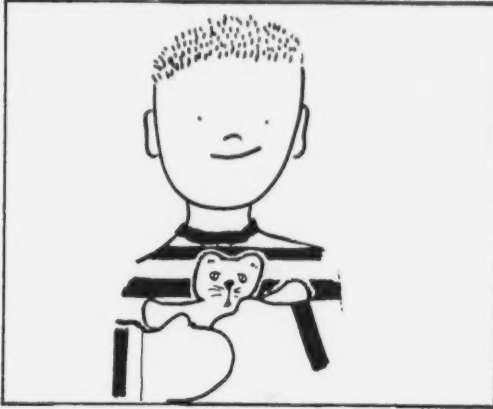
Bereiter & Scardamalia (1982) suggest the amount of text writers this age produce is roughly comparable to what they would produce in one turn in a casual conversation. They contend lack of a conversation partner makes the writer run dry. Encouragement, even a prompt as slight as a questioning 'yes?' has been shown to trigger the writer's search for content anew. If no encouragement whatsoever follows, no more text is generated, either: young writers do not easily continue a search for content themselves. Neither do they evaluate what they come up with while searching, let alone reject material that might be irrelevant, confusing or out of tune. Immature writers do not view their texts thus detachedly. All things are equally welcome, all is put into the text in the exact order it is generated in: 'the less proficient (writers) ... waited until an idea occurred to them ... they seized it and began to write until they were forced to pause ... to search for additional ideas or words to complete the sentence ... they linked discrete ideas and words together with little regard for overall meaning or the anticipated needs of the reader' (Birnbaum 1982: 255). This may, in texts of only a little more length than (2.1) and (2.2) result in helter-skelter prose, posing some problems to the reader. Text (2.3) may illustrate this. It is a text written to describe the drawing of a child.

- (2.3) He has a crew cut.
A bear that looks sullen.
A checked sweater.
And beady eyes.
And an oval face.
He laughs.
His bear has cat's eyes.
(4/23/22/M3; Erwin)

Hij heeft stekeltjes haar.
Een beer die sip kijkt.
Een geruiten trui.
En kraanogen.
En een ovaal gezicht.
Hij lacht.
Zijn beer heeft kattenogen.

Readers in all probability have to put some extra effort in recreating boy and bear described here, paying conscious attention to what comes where. The picture Erwin describes is shown in figure 2.2. By drawing a line in the picture from crew cut to sweater to eyes to face (to mouth?) to the bear's eyes the wanderings of the writer are brought forward pointedly. Text and reader would profit from a reorganization of the text (such as, e.g.: *He has a crew cut. And an oval face. And beady eyes. He laughs. A checked sweater. A bear that looks sullen. His bear has cat's eyes.* Drawing along a line through the picture you will notice the essential difference).

Figure 2.2 Drawing of one of the children testees could choose to describe.



However, rereading in order to evaluate text is something immature writers very emphatically do not do, in all likelihood simply because they cannot do it. Evaluation of texts implies the ability to create a distance, to read with the eyes and the mind of a stranger. Young writers do not distance themselves from their own texts. What is more, it is amply documented that, while reading their prose out loud to Miss and classmates, as almost all young writers love to do, students do not always read what they wrote. They spontaneously insert material on the spot, replacing or adding to the material present, unwittingly recreating the story they wanted to write.

Syntax

Junior prose for the greater part consists of the simple syntactic structures that characterize speech. Notice the more or less identically structured sentences in texts (2.1) to (2.3), structures pivoting around 'be' and 'have'. Information is listed by coordinating chiefly, so in comparison to the total number of words, the amount of information the texts contain is rather scant (Hunt 1965, 1970; O'Donnell et al. 1967; Loban 1963, 1976). Scardamalia et al. (1982) suggest that young writers may have a very limited syntactic repertoire, condemning them to use syntactically identical frames over and over again. It is hard to conceive how this could be true, students this age being grammatically mature language users. It seems rather more likely that novices do not fully exploit their syntactic versatility. In all probability, they are not the only ones: all immature writers seem to stick to a small set of stereotyped sentence structures. In this context findings of experimental research not at all related to writing research seem to be relevant. There, it was found that differing conceptual structures (triggered by a series of slides presented to subjects) tended to be cast in the same syntactic structure (Bock & Loebell 1990), suggesting not only that the conceptual structure of a sentence and its linguistic structure seem to be

separate entities, but also that 'whatever the operations that create sentence frames, they are prone to repeat themselves' (ibid: 29).

These tender aged writers, however, seem to be handicapped extra as, unable to think of any other way of saying things once the intended meaning has been given its syntactic form, they seem 'to be controlled by external forces including their own encoded words' (Birnbaum ibid: 255). Novice writers do not seem to realize that what they 'said' might as well have been said using other words or another structure: 'they feel that the way they wrote was rather inevitable. ... They are attached to their words with a tenacity that is surprising' (Smith 1982: 132). Still, they have been shown to be able to give up this attachment to words when asked to rephrase their writing, giving and taking the odd word and putting in synonyms (Bereiter & Scardamalia 1982). However, they are unable to successfully reframe their sentences. It seems as though the syntactic structure the conceptual structure is framed in is iron clad, as if content once given its form cannot be given any other form.

Whereas in conversational speech the repeated use of a small set of syntactic frames may very well go unnoticed, in writing it does not. Experienced writers know this. They usually put some conscious effort into varying the syntactic patterns they employ³, juggling parts of speech, perhaps deciding not to use another passive and thus thinking over alternative ways to cloak one and the same conceptual structure syntactically. In doing so, they treat language as an object of thought. And, what is more: they know how to handle this object. They manipulate it, they literally got language at their command. They control it. Young writers don't.

Seven-year-olds, thus, cannot easily or successfully alter the syntactic structures they have used. But of course they are very well able to patch up awkward structures. Miss has to call the errant strings to their attention, though, as the students themselves do not spontaneously reread, let alone rewrite what they have written. If they have to, or when they read out their stories to Miss and classmates, they read what it should say, not what it actually says, be this whole passages, single sentences or isolated phrases, corroborating the speculation (see section 2 of this chapter) that writers as young as these seem to be hardly aware of their detecting and correcting erroneous structures. Keen on meaning, they are unhesitatingly recreating the contents of their story literally, unconsciously setting to rights whatever there was wrong with the form it was cast in. Be that as it may, they are able to consider formal aspects of their writing, viz. if Miss has them read out a particular passage word by word. Then, again, they unhesitatingly recognize and repair 'something you do not say like that' (see, e.g. Perera 1984).

In brief, this section on the first attempts of conventional writing of tiroes brought out that, though they know language can be written down, though they

3. And if they don't they expressly do not do so, in order to effect whatever effect.

are aware of some of the essential differences between spoken texts on the one hand and written on the other, their texts are stamped with the traces of the loose, conversational style of spoken language. Very young writers may be considered to make maximum use of the natural human endowment of language competence and of skills learned through ordinary social experience, but their writing is also limited by this 'knowledge telling' (Bereiter & Scardamalia 1987: 4-5). Lack of conscious control seems to be at the heart of the problem. This conscious control develops only later: 'younger children still failed to exert control over (written) language, whether they were generating or evaluating it' (McCutchen 1988: 317). Some hold that, as this is the case, students as young as these simply cannot efficiently be taught to write (Nold 1982: 20). Others state that 'teaching children the tricks of the writing trade does offer them models to follow' (Daiute 1989: 1). Whatever is the case, these very young writers couldn't care less. They are confident writers, not in the least hindered by their utter lack of mastery.

2.4 Developing a monitor for written language

Content and organization

Texts of older but still definitely junior writers tend to be longer than those of the absolute beginners. This seems to be a matter of course: being older, these writers know more and thus got more to tell. But this, it seems, is not at the heart of the matter. What is far more striking is their lessening dependence on conversational strategies when writing. After all, beginning writers have a lot to tell, too, but they don't as they are not prompted to do so. A writer with only a little more experience seems to be able to keep going all by himself, perhaps through mediation of an internalized conversation-like prompting partner (cf. Bereiter & Scardamalia 1982; Daiute 1985). For an example, compare (2.4), a text written by Jonathan when asked to write on his favourite pet, to (2.1), the one written by Walter, his junior by a year.

- (2.4) Salomo
Our pet animal is called Salomo.
It always hops up and down in its cage
We were watching television
(a piece of music) then Salomo started
warbling loudly all of a sudden!
Every time he heard a flute
he started warbling too⁴ and such
fun we and he had
Salomo hopping and warbling
away.
Sometimes he also flew
You see Salomo is a bird.
(7258/7; Jonathan)

Salomo
Onze huisdier heet Salomo.
Het hipt altijd heen en weer in zijn kooi
We waren naar de Tele Visie aan het
kijken (een muziekstuk) toen begon
Salomo op een zo hard te fluiten!
Telkens als hij een fluit hoorden dan
begon hij ook te fluiten en een
plezier dat we en hij hadden
Salomo hipte en floot maar
raak.
Soms vloog hij ook
Maar Salomo is dus een vogel.

4. In Dutch birds and flutes both *fluiten* (Eng: to flute). Hence the 'too' in the text.

So, the texts of such writers differ quantitatively from Benjamin texts. There is a qualitative difference as well, though, as the first traces of attunement to public and goal are to be found in texts of nine- to twelve-year-olds. Their thus deliberately positioning themselves evinces junior writers' ever so hesitatingly distancing themselves away from their text, causing these to be more self-contained. Vocabulary, point of view, structure of separate sentences and of the text as a whole testify to their leaving their ego-centered stance (cf. e.g. Kroll 1981; Perera *ibid.*). For an example, notice the way Floor in (2.5) tries to professionally act out the role of guide when describing a chair.

- (2.5) Here, you see chair with flowers
at the top you see a decorated
cloth for the head. it is a soft
easy chair so a lunch chair*. You do
not see the legs because pieces
are made over them.
It has curved arm rests.
In between the flowers are wisps.
The chair is very wide.
The cloth on top is covered with
flowers.
(3/31/06/M3; Floor)

Wij zien een stoel met bloemen erop
aan de bovenkant ziet u een versierd
kleedje voor je hoofd. het is een zachte
ruststoel dus een fortuin. Je ziet de po-
ten niet omdat er lappen overheen
zijn gemaakt.
Het heeft ronde leuningen.
Tussen de bloemen staan slinteren.
De stoel is heel breed.
Het kleedje bovenaan is bekleed met
bloemen.

Writers this age are not documented to evaluate their texts, however. So, as they still avidly seize all they come up with while searching their memory for content, their writing tends to be a trifle disorganized all the same. Evidently they hover over the text at hand at a height of a few millimeters only.

Syntax

Syntactic structures older students use in writing tend to get longer than those most frequently found in their everyday speech (cf. Hunt *ibid.*; Loban *ibid.*). This lengthening is caused by an increase in complexity, students no longer coordinating but embedding information, thus also expanding content precision and elaborating explicitly (O'Donnell et al. *ibid.*). As a matter of course, target structures pre-eminently are nouns and noun phrases, naming the people, animals and things the text is about and providing ample opportunity to specify their true or fancied qualities. Second in focus are adverbial phrases, specifying the when, how and whereabouts of matters mentioned. Adverbial phrases may also contain nouns and noun phrases which may be modified in their turn. This development, often referred to as 'syntactic maturation', causes writing to be systematically more integrated, or more complexly structured syntactically than

5. In the original, Floor uses *u*, the formal form of *you*, used to address strangers. The first *we* she uses (meaning *we*, translated with *you*) signals she imagines herself addressing more than one person and suggests she uses the plural form of *u* here. This form is especially marked for formality.

6. Floor confuses the French *fauteuil*, in Dutch commonly used for a luxury arm chair or lounge chair, with its phonetic look-alike *fortuin*, meaning treasure.

speech⁷, thus allegedly pointing to a growing 'syntactic fluency' in writing, signifying ease in structuring content and an increasing repertoire (for a discussion of the notions 'syntactic maturation' and 'syntactic fluency', cf. section 6). For an example, compare (2.6a) and (2.6b) and compare (2.7a), (2.7b) and (2.7c). The former two are descriptions of an antimacassar to be found on the drawing of an easy chair, the latter trio describe the elephant on the mug in figure 2.1. Square brackets roughly represent the syntactic structure of the sentences.

- (2.6a) [there is [a cloth] [at the back]]
[er legt [een kledje] [op de leuning]]
(1/31/06/M3; Debbie)
- (2.6b) [[at the top] [you see [a decorated cloth [for the head]]]]
[[aan de bovenkant [ziet u [een versierd kledje [voor je hoofd]]]]
(3/31/06/M3; Floor)
- (2.7a) [there is [an elephant] upon it]
[er staat [een olifant] op]
(4/13/29/M3; Rob)⁸
- (2.7b) [the elephant has [[a rug [on its back]] [and tusks [and a dangling trunk]]]]
[de olifant heeft [[een kleed [op zijn rug]] [en slaganden [en een hangende slurf]]]]
(6/21/18/M3; Jelle)
- (2.7c) [upon it there is [[a large elephant [with a rug [on its back]]] [and a long trunk [and large tusks [and not very big eyes]]]]
[daarboven staat [[een grote olifant [met een maije [op zijn rug]]] [en een lange slurf [en grote slaganden [en niet zulke grote ogen]]]]
(5/21/14/M3; Robert)⁹

Generally speaking, students this age are not the confident writers they were when they were younger. It seems as if they are no longer able to rectify awkward or deviant syntactic structures. Moreover, they have been found prepared to go to any length in order to avoid perfectly fine expressions of everyday speech in their writing (cf. e.g. Perera *ibid.* 1986; Bereiter & Scardamalia 1987). The 'errors' this results in had best be welcomed as signs of growth, to quote Weaver (1982), as they might be argued to bring out that writers, as yet unable to pinpoint, let alone correct problems, grow to be aware of problems at least (Daiute 1985). It has been suggested eight- or nine-year-old writers start 'experimenting' with written language, no longer building solely on their oral language capacity, but also hesitantly testing hypotheses on written language. By thus differentiating between the two languages they grow to be writers (see esp. Kroll *ibid.*; Perera *ibid.*). These 'differentiating' writers thus consciously shift their --heretofore: unconscious-- attention to the writing of

7. As preschoolers 'knew' it to be all along, cf. section 2.

8. Notice the resemblance this structure bears to the one in (2.6a).

9. Notice the first half of this sentence is the spitting image of the one Floor used in (2.6b).

writing, no longer merely 'selecting a meaning to be expressed and selecting a surface form in which to express the meaning' (Kaufner et al. 1986: 138) but also evaluating what is written 'for appropriateness of word choice, meaning, grammatical form etc.' (ibid.). They have been caught struggling to gain the upper hand, guiding themselves during 'the process of constructing a segment of language subvocally, then writing it, constructing another segment, writing it and so on' (Bereiter & Scardamalia 1987: 193). This, in all likelihood, represents 'a cognitively more advanced way of composing' (ibid.). Usage of such subvocalized speech is a phenomenon indicative of a brain performing a complicated or demanding task (Williams 1983). The youngest of writers have hardly ever been found to thus use covert language during pauses in writing. Older, slightly more proficient writers tend 'to reflect over a wider range of alternatives ... The more proficient seemed not only to know how to think but also what to think about' (Birnbaum *ibid.*: 250).

Differentiating oral from written language in all probability is a more or less natural developmental phenomenon in a literate society: 'the child's recognition of the literate properties of text depends upon metalinguistic abilities ... that emerge in middle childhood' (Herriman 1986: 167-8). The five year gap yawning between children's recreating a typical narrative on basis of their tacit knowledge of grammar and genre (cf. section 2) and their supposedly determined creation of a typically written text themselves pointedly illustrates conscious control to be crucial.

However, thus purposely monitoring written language might very well be vexatious to students, as they try to create written language wittingly but lack explicit knowledge. So they might very well detect that something is wrong without knowing what exactly it is that is wrong. Or they may find fault with a passage for being inadequate without being able to pinpoint what it is that is inadequate, let alone correct it. Language users are, by nature, not disposed of knowledge on the foreign language standard writing is. There is nothing that tells them what they need to know: 'is this fine in writing or is it not?'. So, when novices differentiate writing from speech, they may be seen to search, not find, then hypothesize and try out rules, never knowing exactly whether something is wrong, what is wrong, why it is wrong and what to do to correct what feels wrong. All novices, not just the juniors, are so plagued: 'even adults with limited writing ability have been shown to be unable to remedy problems that they can identify in writing. Such information alerts us to the possible partial effects of self-monitoring and suggests that the writer may have successfully identified a problem after self-questioning but does not know the appropriate way to improve the text' (Daiute 1985: 136). Flower et al. (1986) maintain that in order 'to see a problem in the text as a meaningful, familiar pattern' writers need to 'separate the problem from the noise by focusing attention on the features that matter'. To do so they have to be disposed of 'relevant past experience and specialized knowledge', operating on no matter what level of awareness (*ibid.*: 49).

Thus, emergent conscious monitoring capacities suggest explicit rules should be provided to help novice writers out of their predicament, thus enabling them to guess educatedly what is good writing and what is not. Generally, conscious control (or, in the Krashen 1977 notation: Monitoring), driven by explicitly learned rules does not at all ameliorate performance of language users (cf. esp. Krashen 1977, 1978). It does, however, ameliorate performance in writing (cf. section 1); a fine illustration of the queer language writing is.

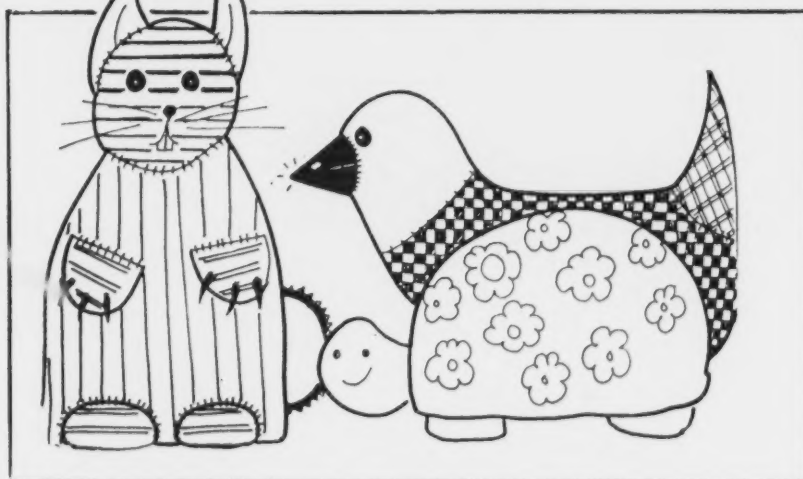
2.5 Zooming in on sentences

Walter, in (2.1), nor Rob, author of (2.2), used punctuation marks. Not using punctuation marks does not mean not knowing they are, or have to be, there: 'Even if writers have never used a mark of punctuation, they are usually familiar with it. Shawn recognized the colon. "My dad uses them in his writing," he says. "But one thing; I can't read his writing".' (Graves 1980: 573). All students, once they are trained to become conventional writers, know sentences start with a capital letter and end with a full stop or a question mark. Some forget about this completely while writing, some very young ones use a capital letter and full stop for every word, some may start their story with a capital letter and end it with a full stop, some seem to think one line of writing is what comes in between a capital and a mark of punctuation (also cf. Edelsky 1983). At a very early age, however, students have been found to abandon 'earlier attempts which interfered with such intact units as clauses, phrases and adverbials ... in favour of honouring syntactic and semantic boundaries' (Cordeiro 1988: 71). This does not mean students go about punctuation conventionally. They might start each new sentence on a new line, use bars to separate sentences, number their sentences or use capital letters and punctuation marks. Moreover, as syntactic and semantic boundaries abound within sentences, segmenting monologues does not mean that sentences are what is segmented. Erwins's text (2.3) bears witness to this, and so does Myra's description of the duck in figure 2.3, presented below in (2.8).

(2.8.) The duck has round eyes.
A pointed nose.
Checked body.
And checked tail.
And the tail is up.
(5579/6; Myra)

De eend heeft ronde ogen.
Een spitsen neus.
Geblokt liggaam.
En geruilen staart.
En de staart is omhoog.

Figure 2.3 Drawing of rag dolls. Testees had to choose one and describe it.



Punctuation such as this again illustrates novice writers build on their oral language capacities and, more in particular, on their conversational skill. Conversational speech does not come in sentences. Rather, 'it is produced in a series of spurts' (Chafe 1986: 15), 'spoken language being typically constructed of a series of independent clauses and clause fragments' one at the time (ibid: 28). Chafe named these spurts intonation units. In his view, intonation units function as basic units in speaking. He takes them to be the linguistic representation of the amount of information a speaker is able to focus on and verbalize at one time (Danielewicz & Chafe 1985: 215). Thus limited by a speaker's memory span, intonation units typically are about five or six words long and take about two seconds to be uttered. In general, all are preceded by a pause and most are smaller than what in writing makes a sentence. These intonation units, syntactic and semantic wholes always, are strongly reminiscent of the phrasal constituents that have been found to be the level complete sentences are organized on (Bock & Loebell *ibid.*). Also, research into pauses occurring in speech as well as in writing brought out that these are typically patterned around structures that are larger than individual words, but smaller than full clauses (Ford & Holmes 1978; Bereiter & Scardamalia 1987; Matsushashi 1981).

It might be a bit rash to conclude that intonation units, boundaries coinciding with those of syntactic and semantic units within the sentence, are key structures in writing. They are, however, obviously target structures in punctuation of juniors purposely producing written language. Examples (2.9) to (2.12) go to show that such writers do not punctuate at random but in all likelihood base their punctuation strategy on these intonation units, partitioning off an after-

thought or two, slicing up full clauses into fragments smaller than what is conventionally called a sentence in writing¹⁰.

- (2.9) She is wearing a sweater. With dots.
Ze heeft een trui aan. Met stippen.
(1/31/00/M0; Leila)
- (2.10) They were dancing and singing. Until it got light again. ... And the next day. Was the girl's birthday.
Ze dansten en zongen. Toidat het licht werd. ... En de volgende dag. Was het meisje jarig.
(4/31/33/M5; Petra)
- (2.11) And when I am riding and when am ready. Then she always may ride it.
And if she is in trouble. She can always lean on me.
En als ik aan het rijden ben en als dan klaar ben. Dan mag zij er altijd op.
En als ze problemen hebt. Dan hebt ze altijd steen van mij.
(72/6.5; Angélique)
- (2.12) When they arrived at school all children came to the two boys. And said you dare you against a pair of rain boots. ... They must have been walking for 1 1/2 hour. In the woods wearing the old boots. They got home at 2. To put on dry clothes.
Toen ze op school waren kwamen alle kinderen naar de twee jongens toe. En zeiden je durft je wel tegen een paar regenlaarzen. ... Ze liepen zeker een 1 1/2 uur. In het bos met de ouwe laarzen aan. Ze waren om 2 uur thuis. Om droge kleren aan te trekken.
(0516/7; Arno)

Large scale assessment studies have brought out that segmenting prose into sentences is no easy matter. Dutch fifteen-year-olds have been shown to be all at sea as far as sentences are concerned (Blok 1987; De Glopper 1988; Van den Bergh & Kuhlemeier 1989). Eighteen-years-old 'basic writers' have been widely reported to be, though older, not much wiser (cf. e.g. Shaughnessy 1977; Bartholomae 1980; Haswell 1988a, 1988b). Zwarts (1990) found that (Dutch) twelve-year-olds had difficulty demarcating separate sentences by full stops, whereas they actually were rather good at segmenting sentences in smaller units by, e.g. commas. This is no surprise: commas come where in speech comes a pause¹¹, and students are skilled markers of what one might, perhaps oversimplifying, call intonation units.

Obviously, then, sentences in writing are a breed of their own: they do not come naturally to language users, once again bringing out the curiosity written language is. A crucial characteristic of the rarefied tribe sentences in writing are seems to be their being 'complete': a sentence in writing is a 'whole' of one or more words (a 'whole' of one word making a sentence is, e.g., 'Stop!'). This whole is difficult but in all probability best to characterize in grammatical terms. Traditionally, grammatical subject and predicate are considered to be the

10. The fragments in (2.9) are taken from a description. Fragments in (2.10) originated from a narrative on 'the shop-window mannequin that wanted to celebrate her birthday'. The fragments in (2.11) were written in an exposé on the writer's best friend, who, as she tells us, is allowed to ride her horse. The fragments in (2.12) come from a narrative on a pair of rain boots that went for a walk.

11. Note that this actually is how students are supposed to get to grips with using commas: 'put a comma where you hear a pause'.

syntactic as well as semantic fundamentals of sentences. In very general terms, the 'subject' of the sentence has a close general relation to 'what is being discussed', the 'theme' of the sentence, with the normal implication that something new (the predicate) is being said about a 'subject' that has already been introduced in an earlier sentence¹² (Quirk & Greenbaum 1973: 11; also cf. ANS, p. 778).

Erroneous punctuation bears evidence to the problems language users confront when they have to segment running written text into sentences. It very emphatically does not mean writers erratically frame syntactic structures. Of course they don't. Building on their tacit knowledge, they simply cannot go wrong there¹³. They go wrong where their tacit knowledge is of no use: in demarcating the fully-fledged sentences of written language.

2.6 Effects of grammar instruction upon writing

Writing is frozen language requiring extensive conscious monitoring. Basically, sentences carry the elaborately structured meaning of a written text. Such sentences do not come naturally to language users. They need to learn about them: what they are and how to use them. Such instruction touches on the structuring of sentences, i.e. on syntax and semantics. So, here, grammar instruction comes in. This section discusses the effects of various grammar treatments meant to beneficially influence writing.

Miscellaneous grammar instruction studies

Braddock et al. (1963) as well as Hillocks (1986) approvingly cite Harris (1962) who stated that the 'study of English grammatical terminology had a negligible or even relatively harmful effect upon the correctness of children's writing' (Braddock et al. *ibid*: 83; Hillocks *ibid*: 133). It seems, then, that whatever traditional grammar instruction does, it does not improve quality of written language. Harris and Stevens et al. (1987) found that grammar instruction is helpful, however, when it is embedded in writing activities of students. Discussing problems with sentence structure, practising solutions, imitating models, making attempts to apply what is learned in free writing and discussing that again is what ameliorated quality. Thus writing, not grammar, came first. Grammar is subservient to writing.

12. Thus pointedly bringing out the natural theme/rheme distinction in discourse.

13. Assessing syntactic quality of essays in terms of T-units will bring this out. A T-unit, or, in full: a minimally terminable unit, measure and name coined by Hunt (1965), is a sentence including all subordinate sentences. In an analysis of writing in terms of T-units, punctuation is ignored (also cf. chapter 5.2 and 8.1).

Harris instructed his teachers only up to a certain extent in the ways they had to pay attention to 'practice in writing and imitating conventional forms and structures' (Braddock et al. *ibid*: 78). Writing lessons basically were 'extensions of usual composition practice'. Teachers were moderately free to add personal detail to this and so they did: 'most teachers gave cohesion and interest to the work by engaging in a variety of longer projects such as the compilation of a diary, a form newspaper ...' (*ibid*: 78). This contrasts sharply to the 'integrated grammar and composition lessons' in the regular textbooks used by the control group, where there was not much time available for such writing activities.

Though in integrated composition and grammar courses writing comes first, immediately followed by explicitly pointed grammar instruction, there are apparently very different flavours when it comes to such integration. A quote from the Stevens study highlights this.

'The [integrated language arts and writing] curriculum comprised a series of teacher-directed lessons which began with instruction on writing complete sentences and moved step-by-step to instruction on writing narratives, descriptions and other genres. Language mechanics were integrated with the writing activities, and were presented as aids to improve writing rather than as separate skills. For example, when students studied modifiers, they then used them in descriptive paragraphs ...
... students worked with teammates to edit and revise their writing. Teammates edited one another's work using 'peer editing forms' emphasizing both the content and the grammatical correctness of the compositions. Then students revised their compositions on the basis of their peers' feedback' (Stevens et al. 1987: 440)

It seems as though for writing skill to benefit from grammar instruction, the latter should be thoroughly embedded in the writing curriculum, thus turning the isolated, declarative grammatical knowledge into integrated, procedural knowledge: a tool that writers may profitably use while monitoring their writing. It seems highly unlikely that grammar instruction thus deeply interwoven into a writing curriculum still looks like anything one would traditionally name 'grammar'. In order to be brought (not: brought down!) to such a subservient level, traditional grammar needs to be qualitatively transformed completely. The effects of this kind of writing and adjoining grammar course contrast sharply with the effects of a regular grammar course, whose explicitly stated rules on parts of speech constitute a body of knowledge that does not transfer to writing as these rules evidently do not bear on writing. But then, of course, neither did the completely integrated grammar 'transfer', for the sake of its being completely integrated.

Elley et al. (1979) found that what goes for traditional grammar goes for transformational grammar as well: no differences were found in any aspects of writing after three years of either traditional or transformational generative grammar instruction.

The findings of Martin (1980) suggest the explicitness of the grammar instruction is crucial as well. He hypothesized that '... knowing how good writing feels will improve students' writing by allowing them to draw on these internalized resources in their own writing' (Martin 1980: 34) and he presented stu-

dents with model passages that were slightly more mature syntactically than the passages the students themselves produced, thus heavily relying on implicit or unobtrusive learning. After the structure of the models had been discussed, students imitated the model twice, writing about something (shown on a slide) that resembled what was written about in the model passage. Performance of his experimental students provided some evidence as to their internalizing 'how good writing felt', but the spur of the models evidently was too weak to result in (statistically significant) gains over the control group in quality of written language.

Sentence combining

Mellon (1969) was the first one to experiment with sentence combining instruction. He did so because he wanted to enrich the student's language environment 'through the provision of occasions for language use not otherwise likely to occur in that environment' (Mellon 1969: 84). For this he provided students with sets of simple, or kernel, sentences plus explicit grammatical cues for combining these sentences into one. The cues served as signals for the transformations that had to be applied. In the early form the Chomskyan roots of sentence combining instruction are easily recognizable. An example (from Mellon, *ibid*: 22)

Problem:

The children must have clearly wondered SOMETHING.

The bombings had orphaned the children.

SOMETHING was humanly possible somehow. (T:wh)

Their conquerors pretended SOMETHING. (T:infin.-T:cap)

Chewing gums and smiles might compensate for the losses. (T:fact)

The losses were heartbreaking.

They had so recently sustained the losses.

The result of the diverse syntactic operations¹⁴ was the following:

The children whom the bombings had orphaned clearly must have wondered how it was humanly possible for their conquerors to pretend that chewing gum and smiles might compensate for the heartbreaking losses they had so recently sustained.

Mellon considers sentence combining instruction to be a special case of language arts programmes which he thinks 'are most successful when they are least structured, content-oriented and self-conscious, when they place their greatest emphasis upon the provision of widely varied occasions for actual use of language by students' (*ibid*: 83). Grammar-related language practice should, according to Mellon, be 'a-rhetorical in nature and not give the appearance of pertaining to students' work in composition' (*ibid*: 20). The sole purpose of 'such pseudo-production activities ... given in a grammatical format' (p. 21) is the student's 'experiencing of mature sentences' (p. 73). Thus, it seems, Mellon tried to boost students' writing performance by specialized grammar instruction

14. 'T' stands for transformation. Also notice that not all operations are mentioned explicitly. Noun phrase embeddings, e.g., are not.

outside the regular writing curriculum. 'Syntactic fluency' of his experimental students ameliorated; the overall quality of their writing did not. He attributed these 'curious results indeed' (ibid: 69) (favouring the group that had studied traditional grammar!) to 'the effect of the teacher variable, that is, the fact that the highly experienced control teacher was a better composition instructor' (ibid: 70)¹⁵.

As O'Hare was convinced that 'students don't like traditional grammar, not only because it is boring, but also because many of them simply cannot do it' (O'Hare 1971: 27) he brought about one major change in Mellon's study before he replicated it and not met half with success: he put aside all grammatical terminology and replaced the cues by practical little helps (sic). Mellon's sentence combining problem in O'Hare's course looked like this:

The children must have wondered SOMETHING
The bombings had orphaned the children. (WHOM)
SOMETHING was humanly possible somehow. (WHY)
Their conquerors pretended SOMETHING. (IT-FOR-TO)
Chewing gum and smiles might compensate for the losses. (THAT)
The losses were heartbreaking.
They had so recently sustained the losses. (WHICH)

Sentence combining is, and it is so deliberately, a far cry from traditional grammar instruction and as far as O'Hare is concerned sentence combining instruction is not grammar instruction at all. Nor, he contends, should it be¹⁶. O'Hare also vehemently rejects Mellon's interpretation of sentence combining as not having anything to do with real writing. Students 'exposed to sentence-building exercises even in an a-rhetorical setting, are in a very real sense being taught writing', says O'Hare and he compares the activity to football players' practice 'in an 'a-game' setting, often with no opposition' (ibid: 68). The dispute has been carried farther: the a-rhetorical setting of sentence combining courses in general, including the ones Mellon and O'Hare experimented with has seriously been called into question, too: 'even in some seemingly 'a-rhetorical' studies the experimental classes extensively discussed and evaluated their combining products and ... these discussions commonly included contextual considerations' (Kerek et al. 1980: 1066).

In all probability, the debate over what to call sentence combining other than sentence combining derives from the mysteriousness of its doings. Time and again the courses turned out to boost students' syntactic 'maturity' or 'fluency'¹⁷ as well as the overall quality of their essays but it has never become quite clear

15. Now what if this very same teacher had worked his wonders in the experimental condition?

16. But then, O'Hare seems to be rather allergic to the term 'grammar instruction'.

17. Allegedly, syntactic fluency is the syntactic manifestation of cognitive growth or maturation. Hunt, Mellon and Strong carefully distinguish these terms; most sentence combining researchers do not.

how come. There is some consensus on sentence combining instruction providing 'focused practice' (Strong 1986: 4; also cf. Mellon's initial wish to enrich the student's language environment 'through the provision of occasions for language use not otherwise likely to occur in that environment', and Martin's models that were slightly more mature syntactically than the passages the students themselves wrote). Or, as O'Hare speculatively put it: 'Perhaps the syntactic manipulative skill the (sentence combining) students had developed, because it entailed a wider set of syntactic alternatives, invited or attracted detail. Perhaps knowing how does help to create what' (O'Hare, *ibid*: 72).

Those concerned with this kind of writing instruction in general would agree that a sentence combining course accelerates at least in part the natural cognitive maturation of students and that this maturing manifests itself on paper in students' increased syntactic fluency. Some seem to cast doubt on the alleged beneficence of this speedy maturation as there seem to be students not treated with a sentence combining course and thus maturing quite naturally that are found to catch up easily with their sentence combining schoolmates after a couple of months already (Hillocks, *ibid*: 145).

Watts' experiment (Watts 1980) did not result in differences in overall syntactic maturity of (low achieving) students. However, in comparison to their own performance and in comparison to post test performance of the control group she found the decrease in mazes (or garbles, or unintelligible strings) in the writing of the sentence combining group to be statistically significant. Apparently, sentence construction instruction helped these students to monitor or: begin monitor, their writing. In this context, Haswell's findings (Haswell 1981) seem to be especially relevant. He found that syntactic maturity of his initially lowest achieving students thrived on a course in sentence combining. In fact, their syntax flourished so enormously its increased maturity was responsible for the (statistically significant) gains of the whole of the sentence combining group. Taking into account Haswell's groups were rather small ($n = \leq 50$), this is a decidedly impressive effect. But then, of course, though 'low achievers' do not always respond positively to treatment, they are, if they do, generally speaking, fine subjects, as they are so far removed from top-notch performance at the onset of treatment.

Sentence combining construction has also been under attack because of the emphasis on producing complex syntactic structures¹⁸. First of all, as Strong points out, rather than combining, uncombining sentences might do a world of good to some students' writing: 'Students should be taught that clarity counts - and that brevity has force' (Strong *ibid*: 22). Next, Smith and Combs (1980) stated that 'we can achieve through well-specified cues in a single week the equivalent of a whole semester of S(entence) C(ombining)-practice' (Smith & Combs *ibid*: 35). They gave freshmen either an overt cue (suggesting the

18. Or, as Hillocks has it: 'the descriptive measure has become an instructional goal' (*ibid*: 148).

audience students wrote to was susceptible to complex sentences) or a covert cue (viz.: a one week course in sentence combining instruction). Their results made them hypothesize that 'SC may not teach the use of the structures, but simply provide the students with a cue, a covert cue, that they should write longer, more complex sentences' (ibid: 19). Thus, they forcibly threaten to take away the ground from under the feet of sentence combining adherents: 'students may not be learning *how* to construct longer, more complex sentences from SC practice, but rather that they *should* do so ... there seems to be a real difference between what the students may do and what they can do if they are given specific instructions' (ibid: 19, 23; emphasis: S&C). Smith & Combs suggest that the concept of audience is crucial: 'students do not have a well-formed concept of audience but they can alter their writing when given highly specific descriptors of an audience' (ibid: 38) and their suggestion is corroborated by Kroll (1985a; 1985b) who found the same for junior students: when these juniors rewrote a story for students being their juniors in their turn, 'students at all of the grade levels tended to reduce the complexity of the original story, thereby demonstrating a basic sensitivity to the needs of a young reader' (Kroll 1985a: 136).

Sentence construction

In 1967 Christensen pronounced that writing of students should be ameliorated by applying classical rhetorical notions. He considered grammar and rhetoric to be complementary, albeit their 'goals and procedures are quite different. Grammar maps out the possible; rhetoric narrows the possible down to the desirable or effective' (Christensen & Christensen 1978: 61).

Analyzing style of writing of authors such as Hemingway and James, Christensen came upon four characteristic aspects of style. These he made into the principles of his generative rhetoric.

The first principle concerns addition of, e.g. qualities, by using adjectives or adverbs, or details, by using adverbs or prepositional phrases. Comparisons can be used to modify information as well. Next comes the direction of modification: details preceding or following nouns, verbs or full clauses. To the opinion of Christensen, not only modification but also variety in its positioning qualifies good writing. Thirdly Christensen noticed professional writing to have more than one level of generality, details being less general, less abstract than the full clause they are embedded in or added to. The final principle concerns the overall structure of texts. Expert writers vary the texture of their writing, alternating texture rich in detail with transparent, thin texture.

Examples of the sentence types Christensen was preoccupied with are given below.

He dipped his hands in the bichloride solution and shook them, a quick shake, fingers down, like the fingers of a pianist above the keys.
(from Sinclair Lewis)

Calico-coated, small-bodied, with delicate legs and pink faces in which their mismatched eyes rolled wild and subdued, they huddled, gaudy motionless and alert, wild as deer, dead as rattlesnakes, quiet as doves.
(from William Faulkner)

Christensen identified the structural levels with numbers, the basic sentence being number 1, modifiers at all levels of generality following suit. The above passages would be numbered as follows.

1. He dipped his hands in the bichloride solution and shook them,
 2. a quick shake,
 3. fingers down,
 4. like the fingers of a pianist above the keys.
2. Calico-coated,
 2. small-bodied,
 3. with delicate legs and pink faces in which their mismatched eyes rolled wild and subdued,
1. they huddled,
 2. gaudy motionless and alert,
 2. wild as deer,
 2. dead as rattlesnakes,
 2. quiet as doves.

The four principles of generative rhetoric constitute Christensen's programme on the 'cumulative sentence'¹⁹. Students start out by observing something they put down in writing by generating a sentence base. Next they need to think of details they want to add: what are these details, where should they go, how prominent are they, what might as well be left out and why? In this way students are trained to consciously construct sentences, to evaluate content and structure of all they want to say and all they want to add, deliberately moulding meaning and form, wittingly creating text. Generative rhetoric, teaching forms that provide slots for detail, thus pregnantly highlights the reciprocal relationship of structure and content. '... the mere form of the (cumulative) sentence generated ideas', says Christensen. 'It serves the needs of both the writer and the reader, the writer by compelling him to examine his thought, the reader by letting him into the writer's thought' (ibid: 28). He pointedly charged the idea that 'we think naturally in primer sentences, progress naturally to compound sentences and must be taught to combine primer sentences into complex sentences -- and that complex sentences are the mark of maturity' (ibid: 26), thus fiercely rebutting sentence combining programmes.

Hardaway's study into the effects of a course in generative rhetoric on freshmen's writing skill (Hardaway 1969) did not show the rhetoric programme to be more successful than the other programme Hardaway experimented with. The study was plagued by seriously disagreeing essay scorers. This may very well be the (major) cause for the lack of experimental effects.

19. And, as a matter of fact, on the paragraph, too.

Brooks (1975) found her twelve-year-olds did not differ as far as syntactic maturity was concerned. Her experimental group, however, did increase its using sentence modifiers²⁰. She did not check the effect on overall essay quality. Faigley (1979) did and he found the overall quality of essays of the generative rhetoric students surpassed quality of the essays of the non-generative rhetoric students. Also, the former's syntactic maturity exceeded that of the latter. However, scores on the renowned syntactic maturity measures were not correlated in any significant way to the high quality of essays. Faigley found that the 'single syntactic variable strongly linked to overall quality (was) the percentage of T-units with final nonrestrictive modifiers' (Kinghorn et al. 1981: 33). Faigley tentatively explains this finding thus: 'The compositions of younger children are sometimes characterized as a series of unelaborated propositions; similar characteristics in the essays of college students reveal a lack of development. Generative rhetoric addresses this problem by stressing the addition of specific details to abstract statements as a means of generating content' (Faigley *ibid*: 204). One might perhaps, slightly speculatively, conclude that structures thus demanding content and content thus demanding structure results in a text that tells. The amount of contemplating content and pondering over plans, however local, pertaining to sentences only, results in carefully elaborated discourse. The quality of such elaborated discourse outstrips that of non-elaborated discourse, however complexly structured (also cf. Hillocks *ibid*: 75-6; 147). Complexly structured sentences are complexly structured sentences. Elaborate discourse is writing.

2.7 Conclusion

At the heart of this study into the effects of grammar instruction upon junior writing is the conviction that grammar instruction at school may very well alert students to the main features of the foreign language they need to familiarize themselves with: standard written language. The most fundamental feature is the sentence. So, says this study, learning what in writing is a sentence need not at all be done in a writing programme. What is more, for a grammar course dealing with sentences to be successful, i.e. for its effects to manifest themselves in writing, students do not have to be engaged in actual writing at all. The expert knowledge on sentences will transfer spontaneously to writing: students want it.

The pivot of this chapter was the increasing control students exert over their doings while writing, brought to the fore by their leaving conversational schemata and by their determined effort to differentiate writing from speech. The

20. Sentence modifiers lengthen mean T-unit length (MTL) and increase clauses per T-unit (C/T), MTL and C/T being two of the three measures that determine syntactic maturity. Evidently Brooks' students did not use enough sentence modifiers; this all the more seems to hammer home Hillocks' hint that in sentence construction instruction the descriptive measure has become the instructional goal.

youngest of young writers write down conversational speech in spite of their knowledge on written language. Writers as young or immature as these do not seem to be able to put to use the knowledge they hold on written language. All of their writing demonstrates an utter lack of control. Nine-year-olds differentiate writing from speech. Symptoms are their planning writing, brought out by the increase in covert language behaviour during pauses in writing. Also, they try to manipulate their own position as a writer versus that of their public. Moreover, their producing concise writing more frequently reinforces the observation that these writers put conscious effort in creating writing. The same goes for the greater variety of syntactic frames they employ.

Symptoms may be captured under the heading of conscious control. However frail be this control, students executing such control over what, how and for whom they are writing evidently are able to exploit all knowledge on the matter they possess. This, of course, is precisely what conscious control means: a conscious search for knowledge that, once found, is conscientiously used to build hypotheses to be tested.

Searching through their knowledge, however, students will be unable to find anything useful on this dinosaur among all languages: standard writing. There is no body of knowledge available. This study maintains that grammar instruction has to provide writers with such knowledge. The conjecture is that learning about sentences, thus building a new domain of knowledge pertaining to writing, poses no problems to student writers gaining control, as these are so evidently able to consciously consult their knowledge and act upon it. So, this sort of explicit knowledge on written language, or, more specifically: on sentences in writing, will transfer to writing.

What should such grammar instruction look like? Obviously, it should not be of traditional or transformational flavour. Numerous studies have brought out that neither traditional or transformational grammatical knowledge ameliorates writing. If there is any truth in the speculation that writers need to know explicitly what distinguishes oral from written language, the lack of success of courses such as these is no surprise: they do not all concern this matter. Programmes that do concern this matter are courses that provide grammar instruction where and when it pertains to writing and courses in sentence combining or sentence construction.

Students seem to be able to experiment successfully with recently introduced grammatical *bric-à-brac* in writing, but then, integrating grammar instruction this way in the writing curriculum impedes studying transfer.

The wonders that sentence combining courses worked are surrounded by rather prominent, intriguing issues that bear on its effectiveness. Or rather: bore, as sentence combining is nothing it was when it was born. The issues lost nothing of their intrigue, though. Is the syntactic maturity, or fluency, courses in sentence combining so generally seem to kindle no more than the result of a mere cue signalling desired complexity of structures, especially to low achieving students? Does it do anything more than provide students with a set of sentence

frames and a series of devices to be exploited successfully in writing in order to heighten sentence complexity?

The section on generative rhetoric suggests the constructing of sentences Christensen propagates is a superior writing programme because it draws attention to structures' dependency on content and vice versa. Contrary to traditional or transformational grammar, no matter in what guise they come, generative rhetoric provides a minimum amount of explicit rules. What rules there are help writers mould meaning while thinking over their writing.

Generative rhetoric highlights the basic building blocks of writers: the sentence. Neither traditional nor transformational nor integrated grammar do so. Sentence combining does. But whereas sentence combining pointedly emphasizes the need for complexity, cueing students to optimally exploit their natural linguistic endowment in producing intricately interwoven prose, generative rhetoric, guiding sentence plans, drawing attention to slots for elaboration, thus generating content operates on the essential planning and writing unit writers need to know about: the sentence. This, then, seems to be the sort of instruction called for: a grammar course that deals with sentences mainly, explaining what these are made of and how they expand.

COPING WITH THE NUISANCE OF ACTUAL PRACTICE

*'...So remember: a greengrocer's assistant.'
'Do I really look like one?' asked the Doctor, trying to keep
the small cap from slipping off the back of his head.
'Well - you do,' said Matthew. 'Turn your coat-collar up and
don't talk too grammatic - no fancy lingo - see?'*

Hugh Lofting, Doctor Doolittle's caravan

This chapter deals with all sorts of matters pertaining to the design of the experiment undertaken to study the effects of grammar instruction upon junior writing. It orbits around crucial issues: what instruction is hypothesized to bring about what changes and how to engage safely in such a venture. Sections 1 and 2 will deal decidedly sketchily with the former matters, i.e. with treatment and dependent variables; chapters 4 and 5 will amply deal with these. Emphasis here, then, is on the design of the study and on all measures taken to safeguard it from harm. All pass in Indian file in section 3 to 7. Section 8 summarizes the lot.

3.1 What instruction ...

Chapter 2 wound up by stating that a course in generative rhetoric seemed to be the most likely candidate to help junior writers where they actually need help: in recognizing and exploiting the fully blown sentences of written language. This is not to say that the Christensen programme will, or can, or should, be transferred bluntly to Dutch classrooms, albeit these are filled to the brim with those eager nine- and ten-year-olds, pens poised. In its genuine guise, the Christensen programme comprised a thorough drill in grammar as well as a good deal of practice in writing. Such a course, whatever its beneficial effects, is not called for in this study as here instruction in merely grammar is assumed to be precisely the booster students need. What this study calls for is a downright grammar course that deals with sentences in writing: what these are made of and how they expand. The context is to be strictly a-rhetorical by nature. If the course has students write, it should have them write (or: construct) separate

sentences and sentences only. Its effect is to be contrasted to a traditional course in school grammar and to a course that has students write and write only. Both are interesting competitors for the grammar of sentences. Those working in Dutch primary education foster a deeply rooted belief in the beneficial effects upon writing of the sort of grammar instruction traditionally found in student's textbooks. As, according to the advocates of this kind of instruction, writing benefits from naming parts of speech and parsing sentences, these are justifiable ends in themselves. Obdurate opponents of this kind of grammar instruction will have it that as practise makes perfect, practise is what students should do. Thus, they train their students directly in writing. So, the three courses to be contrasted are:

- Direct Writing;
- School Grammar;
- Sentence Construction.

Chapter 4 concerns construction, try out and revision of the three programmes.

3.2 ... and what changes.

As put forward in chapter 2, writers about eight to ten years old start monitoring their writing, consciously differentiating written from oral language. As they possess no specialized knowledge on written language, though, they guess in ignorance as to what it is that turns writing into writing. Contrary to the guesses any language learner naturally makes in first language learning, educated guesses on the written variety of language are impossible to make. A case in point is what student writers take to be a sentence. Junior writers should be explicitly instructed in what distinguishes writing from speech. As this study has it, students so badly need to know this, they will use this declarative knowledge even if it is taught outside the writing curriculum. That is: the explicit knowledge on written language students obtain will transfer to writing. This will result in better writing.

This study professes that students do not need to be taught what are or are not syntactically well-formed strings: their writing fluency is fine and no treatment will or can affect it. Novice writers should, however, be taught how these syntactic strings have to be grouped into sentences. As sentences, i.e. the fully fledged sentences of written language, do not come naturally, language users should, first and foremost, know what in writing constitutes a well-formed sentence or what elements must be present in a sentence at least for it to be complete. Correct punctuation of these sentences is to be taught in tandem. It is hypothesized that a course in writing (i.e.: Direct Writing) nor a course in traditional school grammar (i.e. School Grammar) provide such information. Hence they will yield no positive results in sentence awareness, even though the traditional school grammar course successfully teaches academic linguistic knowledge. A grammar of sentences (i.e. Sentence Construction) will yield positive

results, as this brings about explicit knowledge on the grammatical peculiarities of written language.

Also, this study hypothesizes that this explicit knowledge on grammatical peculiarities of written language entails recognizing and exploiting sentential nooks and crannies that information can be put into, resulting in typically concise writing. Thus the Sentence Construction course will also beneficially influence conciseness of students. Neither of the other courses will. As, finally, students taught what are sentences start to compose meaningful sentences consciously, using sentences for the scaffold they are, their prose will turn out to be superior in quality to prose written by students treated by either of the other courses.

So, there are five variables and five hypotheses. These are:

- writing fluency,
hypothesized to be fine for all students; no treatment will affect it;
- academic linguistic knowledge,
to be affected equally beneficently by Sentence Construction and School Grammar and not by Direct Writing at all;
- sentence awareness,
hypothesized to be ameliorated by the Sentence Construction course only;
- conciseness,
another variable that Sentence Construction students are hypothesized to perform best on;
- overall writing quality,
the final variable that Sentence Construction students will perform best on.

Chapter 5 reports on the construction, try out and revision of the measuring instruments for these five dependent variables.

Now, what if Sentence Construction students are to be taught how to construct texts immediately after? And what if Direct Writing students are to be taught how to construct sentences immediately after? Would sentence awareness of students treated either way turn out to be equally affected? Would conciseness and overall writing quality in the end amount to the same? Or would all that bore on sentences be lost and forgotten the moment instruction by Sentence Construction would stop?

And is it not likely that traditional school grammar die-hards will bring forward that in this study the School Grammar programme is a born loser; that the course comes to nothing because students do not write and thus are withheld practice?

These issues are not at the heart of the study; they are explored all the same¹.

1. As a matter of fact, it is a mere freak of fortune the study allowed these explorations to be indulged in at all (cf. chapter 6.1). Also, thinking over questions such as these, one is easily carried away into asking more and more and more of these questions. As these cannot be answered anyway, I'll simply leave these for what they are.

Concerning the first set of questions, it is hypothesized that sentence awareness of students and conciseness and overall quality of their writings will thrive most if students are treated by the Sentence Construction course first and taught how to construct texts only after that. That is to say, Sentence Construction students will not forget about sentences and, secondly, as the Sentence Construction course provides such a sound basis for confident writing, overall quality of writing will increase even more: the Direct Writing course is sown in fecund soil.

Teaching courses the other way round seems to be rather like putting the cart before the horse, then. Though the course in writing might initially affect overall writing quality marginally², the major effect will be brought about by the Sentence Construction course that is to follow.

The question pertaining to the lack of effect of the School Grammar course will be explored by contrasting the effects of School Grammar followed by Direct Writing to the effects of the Sentence Construction course followed by Direct Writing. Hypothesis is, of course, that the former ordering of courses will come to nothing whereas the latter works wonders.

3.3 Design of the study

The study into the effects of grammar instruction upon aspects of writing skill of primary school students actually comprised two experiments: the experiment proper and an extended experiment, meant to explore effects of ordering of specific treatments. The experiment proper always is first in focus. Its design can be described loosely as a randomized pre to post test control group design. At each of the collaborating schools (six in total) nine-year-old students (in The Netherlands: sixth group) were randomly assigned to one of three experimental treatments. In this sense, the study included a true experiment, as by randomly assigning students, not intact classes, to experimental groups, school effects are controlled for: there was no contamination of school and treatment.

Three experimenters, not class teachers, were responsible for treatments. There was not just one experimenter solely responsible for just one treatment: all three worked with all three of the experimental courses. So, neither was treatment contaminated with experimenters. The design is depicted schematically in table 3.1.

2. Notice that this marginal bit of good the course in Direct Writing might do is not one of the hypotheses to be put to the test in the main part of the study.

Table 3.1. Assignment of experimenters (A, B, C) to each of three experimental treatments (X, Y, Z) at the six schools involved

Schools		1	2	3	4	5	6
Treatments	X	A	A	C	C	B	B
	Y	B	B	A	A	C	C
	Z	C	C	B	B	A	A

Ideally, such a randomized block design rules out the possibility that results, whatever they are, are the effect of the experimenter and not of the treatment (so-called experimenter effects; memento Mellon, chapter 2.6)³. Unfortunately, however, the intense seven months training of the experimenters prohibited the experiment to be carried out completely double blindly. Inevitably the experimenters at last were able to work out what the investigation was aimed at. The design itself, then, does not eliminate one of the major threats in experimental research, teacher bias. However, consult section 3 for the measures taken to control teacher bias and chapter 6.2 for their results.

Performance of students was assessed by pre, mid and post tests.

In the extended experiment, follow-up treatments were given to all students, results of which were assessed by a mid and a post test again.

Though the experimental design of the study thus elegantly tames some lethal monsters lying in ambush to butcher research (and researchers) this still does not guarantee that all results experimenting yields are experimental results indeed, brought about by treatment and treatment only. Effects could still be caused by matters that should not have interfered at all. It is not difficult to come up with some of the things that threaten experimenting in real-life, real-class settings. Expert though she is, Miss is human. She may very well not do quite what she is supposed to do, working her head off for one or two treatments only, or skipping crucial matters she thinks are irrelevant after all, mixing up courses and their respective terminology, or dealing with subject matter rather resourcefully but quite erratically. This might make or break treatments on false grounds. Though in real life things might be different, in an experiment testing the effects of instruction type, the teacher is not what should cause effects⁴. By checking protocol adherence teacher bias was controlled for. Second, the amount of time actually spent on subject matter might have differed systematically between treatments. In one or two of them, time thus spent might have been far too little. This may be due to organizational problems or due to

3. Experimenter effects still might pop up as interaction effects of school x treatment. In this design, however, these are torn apart from main effects.

4. Not everyone will agree on this. Strong (1986), for example, remarks on one of the studies he discusses that, while it 'may not be elegant in methodology or in statistical analysis, it shows convincingly the differences that one hard-working teacher can make. I want to emphasize that Argall did not expect S(entence) C(ombining) to do the teaching. She reserved that task for herself' (ibid: 10).

lack of students' motivation. Whatever it is due to, it makes or breaks treatments on false grounds. Time on task observations will disclose whether experimental effects are thus obscured or, worse, nullified.

Third, students might not have done enough work or they may not have done it well enough. This might mean they simply have not acquired all subject matter they were supposed to. Under such circumstances poor performance does not result from the treatment given. One might even wonder whether in such cases there actually is any 'treatment given'. By keeping track of students' progress this will be checked.

Lastly, the regular class teacher might have given students instruction interfering with experimental instruction, thus obliterating or augmenting experimental effects. Hence, the fourth source of evil to keep under control is the regular curricular language arts instruction given to students by the class teacher. In short, this quartet are

- protocol adherence;
- time on task;
- students' progress;
- regular curriculum.

Sections 4 to 7 report on construction of the instruments to control for these nuisance variables. As a pilot study has been executed for only one of these, information on the collection of data in this pilot study will be given in the relevant section.

3.4 Protocol adherence

Though experimenters were allocated to treatments systematically, teacher bias still threatened the results of this experiment as it was impossible to carry it out completely double blindly. Protocols accompanying the courses had to control performance of experimenters and their bias towards one of the three treatments. Performance of experimenters was layed down tightly by these protocols. Moreover, experimenters were trained in working with each of the three experimental courses. Nevertheless, data were gathered on their behaviour (and thus, indirectly, on their bias). To establish their adherence to the protocol it was decided to register all lessons by audiotapes. Per protocol per experimenter four of these (one randomly chosen out of every six or seven) were to be analyzed. Tapes were to be compared to the protocols by ticking correspondences, digressions, omissions and additions. As all lessons in the protocols had separate sections, each subdivided by diverse steps and each culminating in stereotype activities, and as each protocol had its almost ritual, standardized phrases for characteristic course components, comparison of actual and prescribed experimenters' behaviour did not promise to be very intricate. As the

harmfulness of all trespasses had to be judged, the delightful job of listening to all tapes could not be farmed out⁵.

Whether all this was at all possible has never been a point of serious consideration. However, during the pilot study the protocols were used to observe experimenter behaviour and this brought out that it was indeed possible to use the protocols for such a comparison.

3.5 Time on task

As a systematic difference in the amount of time spent on activities pertaining to each treatment, whether due to organizational problems or due to lack of students' motivation, might make or break the effect of a treatment on false grounds, this variable had to be checked. So, data on students' time spent on task were gathered.

To evaluate whether all seemingly relevant activities, or tasks, were included a draft form developed to register time on task of students was tried out for several weeks in the period the experiment ran on trial. Next, when tasks were clustered in categories neatly, this form was used to observe students in all pilot study groups for a total of ten hours in five weeks by two observers working individually. In this way, observers' agreements could be checked. Though observers tended to fulminate against the user unfriendliness of the form, their agreement was satisfactory. Some cosmetic details were changed to appease future observers.

3.6 Students' progress

Poor performance of students might very well result from their not having done enough work in a course, or from not having done work well enough, meaning they simply have not acquired all subject matter they were supposed to. Keeping track of students' progress is a means to control this variable.

First, the amount of covered and mastered subject material had to be checked. Next, systematic differences between treatments in the amount of course material covered and mastered by students had to be investigated. As experimenters had to keep track of students' progress anyway, they were the ones that collected these data. They did so by putting hyphens or ticks on prefab forms for the School Grammar course and the course on Sentence Construction only. For the Direct Writing students, presence was noted as well as the amount of finished work. The character of the Direct Writing course rendered it unnecessary to keep track of students' progress in more detail.

5. As a matter of fact it turned out to be one of the most delightful jobs in the whole project.

3.7 Regular curriculum

Not all of the regular language arts program of the schools involved in the study threatened to interfere with the experimental courses. Some of it would, though it should not. So, all possibly interfering subject material was put under taboo. This was done by first making an inventory of all language arts materials used by the schools involved in the experiment. Those materials that coincided with one or the other experimental course (e.g. *Spelenderwijs Stellen*, the basis for the Direct Writing course, or *Grammatica in Balans*, that inspired the Sentence Construction course) were banned completely for the year of the experiment. Those materials that covered only parts (e.g. *Jouw Taal Mijn Taal*, that the course in school grammar was distilled from) and those that covered similar subject matter were screened closely. All texts or exercises that would, however slightly, interfere with subject matter touched upon in one of the experimental courses were tabooed by simply but very noticeably crossing these out. For this, so-called shadow books were bought: extra copies of the material to be used by the class teacher for the duration of the experiment. The books alerted the teacher to what was taboo and what was not. Moreover, the teacher was supposed to write down in these shadow books the day that non-taboo material was studied in class. If necessary, details as to what the class had been working on and in what way could be supplied as well. In order to regularly check the teacher's accuracy for each teacher there were not just one but two extra copies of the shadow book. Every now and again experimenters were instructed to collect the one in exchange for the other.

3.8 Summary

This chapter concerned some vital issues that come with preparing an experiment. It touched upon the treatments and the differential effects these are hypothesized to bring about in five variables. The bulk of the chapter, though, dealt with matters of control. For, if experimenting yields effects, it should be quite clear what these are effects of: treatment or something else. If experimenting yields no effects, it should be evident why it does not: are treatments a failure or what? So, all that might possibly interfere with treatments and their effects, all sorts of things that might very well affect students' performance beside the experimental treatment have to be kept at bay or should at least be kept under close surveillance.

The experimental design of this study rules out school effects and experimenter effects. Teacher bias and the correctness of instruction is controlled for by audiotaping all lessons and by comparing a random sample of these to the instructions in the protocol. In order to establish whether there were systematic differences in time students spent on actual course activities, time on task observations are to be carried out. The regular experimenter's administration will yield data on systematic differences between students' progress in treat-

ments. The regular language arts curricular activities that may interfere with experimental treatments are banned in shadow books. These shadow books, to be kept by class teachers, also help to check covered subject matter.

ON COMPOSING, CLASSIFYING AND CONSTRUCTING:

The three experimental courses

'I hate grammar. What's the use of it?'

'To teach you to speak and write correctly, so that you can be understood,' said Mrs Garth, with severe precision. 'Should you like to speak as old Job does?'

'Yes,' said Ben, stoutly; 'it's funnier. He says, "Yo goo" - that's just as good as "You go".'

'But he says, "A ship's in the garden", instead of "a sheep";'

said Letty, with an air of superiority. 'You might think he meant a ship of the sea.'

'No, you mightn't, if you weren't silly,' said Ben. 'How could a ship of the sea come there?'

George Eliot, Middlemarch.

In the study into the effects of grammar instruction upon writing of primary school students, effects of three courses are analyzed. One of these actually is no grammar course at all, but the ultimate embodiment of the profound and radical treatment that turns novice writers into apprentice composers. It is called Direct Writing and its makings will be discussed in section 2. Both other courses are grammar courses all right. One is a traditional course clad in a regular modern outfit: School Grammar. It is presented in section 3. The other course is a grammar of sentences: Sentence Construction. It will be amply dealt with in section 4. First, however, this chapter provides a bit of general information on the try out of these programmes. Selection and training of experimenters will be gone into in section 5. Section 6 concludes the chapter.

4.1 Pilot study

To try out the experimental courses and to train the experimenters, the experiment was pre-run on a primary school. Just as if this pre-run was the real thing, students were randomly assigned to one of the three treatments. The form involved in the pilot study (nine-year-olds, in The Netherlands: sixth group) was small: eight students. Experimental groups, thus, were very small. There were two lessons each week during 14 weeks. Each lesson lasted an hour. Due to the school schedule, classes had to be taught on one morning. There was a break

of three quarters of an hour in between classes. During this break the students played outside and, next, were read to by a member of the research team for about half an hour.

Experimenters worked with all three programmes, just like they would have to in the real experiment. Working with all treatments in the pilot study also meant the experimenters had to work with different groups of students. Small though these were, they were expected to (and indeed did) present the experimenters with a whole range of problems concerning daily sorrows such as class discipline, unruly students, students' lack of motivation, keeping up tempo, minding the schedule and so on and so forth.

Detailed manuals for experimenters accompanied courses. In these manuals, or protocols, activities of the experimenters were laid down imperatively. Texts to go with the activities were supplied as well. Experimenters did not have to learn these texts by heart, as they were not supposed to literally parrot the protocol.

In the first half of the pilot study experimenters were observed painstakingly so as to see whether they could, in whatever respect, cope with courses, protocols and students. All lessons were evaluated individually with experimenters in weekly two-hour meetings. All kinds of problems individual experimenters had encountered with students, protocols and treatments were discussed. Potential pitfalls in lessons to come were discussed as well.

The second half of the pilot study was used to try out one of the instruments to control nuisance variables (see chapter 3.5).

4.2 Direct Writing

At the time the experimental courses had to be composed, no separate writing course of the sort wished for in this research was available for nine-year-olds. However, there was material meant to be used every now and then by individual students. This was a sound basis for a real course in writing. The material, *Spelenderwijs Stellen* ('Writing without effort') (Van der Geest & Swüste 1974), consisted of a series of 75 cards, each carrying one writing assignment. Cards were selected in such a way that all kinds of techniques (such as brainstorming individually, shortening a passage, making one long story of three short stories, writing paragraphs, writing with a partner, writing direct dialogue) and all genres (such as narratives, expositions, plays, letters, notices) were included in the course to be made. Writing assignments usually were copied literally in the student's course book (called *Workbook Writing* (Du: *Werkboek Stellen*)). Unlike *Spelenderwijs Stellen*, *Workbook Writing* was organised in seven sections, each with four separate assignments. Also unlike *Spelenderwijs Stellen*, students could no longer choose their own topic and write. All students, albeit each on their own, devoted their attention to one and the same assignment.

Students always made one draft before writing the essay in their workbook. They were told first drafts simply never are good enough and that even a 'real writer' had to think things over twice or even more often than that¹.

Spelenderwijs Stellen had students working on their own. The teacher did not interfere. In *Workbook Writing* the teacher (i.e. the experimenter) introduced the assignments and individually talked essays over with students, but she did not interfere so as to make students change parts of essays they did not, for whatever reason, want to change.

Typical questions of the Direct Writing teacher would be:

- I do not understand this. Can't you write something more about it?
- This interests me very much! Why don't you try and write something more about it?
- This is not how to spell 'stormsgaade' ('storm dammadz'). This is the way you spell it: 'stormschade' ('storm damage').
- Could you read this sentence out loud to me?

Evidently, Miss tried to prompt writers to write more. She did however not tell them what to write or how to put things.

Spelling errors were noticed only. The teacher wrote the correctly spelled word with a black pencil in the right hand bottom corner of the student's draft, so that the writer could spell it correctly when taking the word down in the workbook.

The last question (Could you read this sentence out loud to me?) was meant to alert students to a deviant structure. While reading the corpus delicti, students might automatically patch it up. In that case, Miss told them this was not what they had actually written down and would they change it, please. When students did not automatically alter the sentence, they were asked whether this was the way to put things. If it was, ok, if it wasn't, what was, then? Whatever the students came up with, was ok.

Students enjoyed the first version *Workbook Writing* very much. Experimenters enjoyed it a lot less, mostly because they felt let down by the protocol. It gave them far too few cues on how to start a fine brainstorm and how to keep it going. Also, they thought they could do with some more information on what were supposed to be essential elements of a good Direct Writing lesson. These, obviously, were things that had to be changed for the better.

Also, the experimenters' role was felt to be still too central. Students were found to rewrite part of their essays if Miss apparently wanted them to do so. They did not seem to know why they had to rewrite (or, rather: add material to their essay). Besides, having Miss individually evaluating all essays used up

1. Not surprisingly, students found it very hard to take this in. In the actual experiment, one experimenter got so fed up with students' 'unwillingness' to rewrite their essays she decided to write an essay along with one of her groups. When it came to rewriting essays, after discussing some of them, she pulled out all the stops and started making changes like mad. Her students were utterly perplexed. How could Miss, who knew it all, do such a thing? Still, it made them rewrite too, for if Miss' essay needed rewriting, then how on earth could theirs ...

tons of time. With the mini-group involved in the pilot study ($n=2$) this did not present a great deal of problems, but it certainly would in groups of ten students or more (and probably in less 'large' groups as well).

Peer evaluation was opted for, as this stone killed both birds. Talking over two or three essays with all students need not take much more than 15 to 20 minutes and students, by evaluating other essays, might get an idea as to how to change parts of their own writing. Four such 'peer' lessons were prepared for a small scale try out in a large group (nine- plus ten-year-olds) at the school where the pilot study had been executed. Things worked out fine and so this version of peer evaluation was woven into the Direct Writing course.

Some changes concerning the actual assignments were made, too. Fourteen of these were more or less thematically clustered in three groups. The other groups were formed by twelve assignments that concentrated on all kinds of techniques. All in all, there were 26 assignments.

Organization and purpose of texts are highlighted in the writing course. Both subjects are dealt with in relation to various topics and in a wide variety of ways, starting with simple tasks, steadily working towards greater complexity. In the first task concerning organization, for example, students sort out the chronological order of three out of six drawings, all pertaining to a topic sentence. These are the topic sentences.

There was a storm last night.
When I got out of bed I saw a big branch had snapped off a tree.
The radio said trees had fallen down everywhere.

With these drawings and topic sentences as guides, they next integrate some five to ten sentences they previously wrote, starting with one of the same three topic sentences, into a longer story.

One of the more complicated tasks is to supply information on a fantasy beast in an encyclopedian manner. Prewriting activities for this one are fantasizing on a drawing of a fantasy beast, talking over fantasy beasts in literature (in Dutch: the *kurketrekvogel* (corkscrewbird) in *Hoebelen met Daan* and the *krullevaar* (curly stork) in *Pluk van de Petteflet*), reading (and singing) Carroll-esk nonsense poems on nonsense beasts (*De Bozbezbozzel*, *De Blauwbilgorgel* and *De Ballade van de Pantippel* by Cees Buddingh²), making up fantasy names such as *elidile* or *sticklephant* and leafing through an encyclopedia on animals. Another is to report a story that promises to be like any other but suddenly changes tack. Students tune up to this one by pondering over missing final cartoons in comics and talking over unsuspected silliness of situations.

Purpose of texts is dealt with in relation to preciseness, conciseness and to-the-pointness, mostly. Students prepare would-be pieces² for the club bulletin or their local newspaper, giving information on specific topics or asking for special

2. Talking of purpose ...

gear, they respond to a rather bloody newspaper article on a dog that gets killed after attacking a woman and to crown it all, they actually wrote a letter to a fun-fair or amusement park of their own liking, to ask for information. This assignment, the last of all, got to be a popular one as the fairs and parks almost without exception sent material, some of which was really beautiful. Specialties such as writing dialogues, brainstorming and planning individually and writing in twos and threes was attended to first in isolation. When the course was well under way these specialties got woven into other assignments culminating into, for example, an assignment to plan a story on a long journey carefully in cooperation with a partner and divide the sections each had to draft. After discussing the story that resulted, each student turned the separately written sections into one story. This meant they had to (and did) revise and edit elaborately. When finished, ex-partners changed versions and more often than not were surprised they could hardly recognize their own contribution.

Students got some unobtrusive training in peer evaluation. This started out by having students mentioning the one aspect they liked best about the essay just read to them. Within three or four lessons students were prompted to advise each other on just one thing that might be changed for the better in essays. Guided by a lead on the blackboard, they ultimately kept in mind several pros and cons of the essay read to them to come up with during the discussion. Writers themselves always had the last word.

Lessons typically started off with Miss showing a drawing or telling a story to introduce a topic. This was discussed for some five to ten minutes. Next, Miss would introduce the task in the workbook students had to work on. Details of the task were discussed, too. Then students would prepare a draft to be discussed in the group. The lesson would end with students writing a revised version of their assignment in the workbook.

4.3 School Grammar

For nine-year-olds there are no separate courses available in traditional grammar³. Traditional grammar is touched upon in regular language arts courses, though. *Jouw Taal Mijn Taal* (*Your Language My Language*) (Van Berken et al. 1978) is the course that traditional grammar is touched upon most completely and well-balanced (Van Dort-Slijper 1984; also: Naber et al., 1986⁴). So, it was decided to base the course in traditional grammar to be used in the experiment upon material from *Jouw Taal Mijn Taal*.

3. Lucky devils!

4. For years on end the course is used by a steady 20% of schools in Holland (Bos & Oostdam 1985; Zwarts 1990).

Lessons in *Jouw Taal Mijn Taal* each have one central theme, introduced by a drawing-to-talk-about (Du: 'praatplaat') and a text. All exercises in one lesson relate to the theme. Of the material meant for nine-year-olds (the sixth group in The Netherlands) the text plus some questions bearing on it were included in the School Grammar course. Furthermore, the School Grammar course comprised all *Jouw Taal Mijn Taal* material that touched upon parts of speech (verbs, nouns, adjectives, conjunctions and possessives) and sentence elements (in this course subject, finite verb⁵ and the so-called kernel of the sentence (Du: 'zinskern')). Subjects such as concord, lengthening or shortening of sentences, constructing sentences according to a set structure, punctuation, morphology etc. were also included.

Some ten to twenty minutes of each hour were devoted to the introduction or repetition of subjects. Next, students individually studied exercises in their textbook and wrote down answers in a separate notebook. Usually, at the end of each hour, some ten to fifteen minutes were left to go into another subject more playfully, e.g. by reading and making poems.

Typical questions of the School Grammar teacher are

- In the sentence *John Fox lives in the suburbs of our town* what are the three sentence elements?
- (in the same sentence) What elements can be left out?
- (in the same sentence) What is the kernel of the sentence?
- *Is werkt* (Eng: the finite form *works* of the verb *to work*) a word that goes with *de* or *het* (both meaning *the*)?

The questions are none of them nonsensical. The answers to be given in a course such as this are, though. The three elements of the sentence *John Fox lives in the suburbs of our town* are presumed to be *John Fox lives*, *in the suburbs* and *of our town* (instead of *John Fox*, *lives* and *in the suburbs of our town*). The elements to be left out are not none but *in the suburbs* and *of our town*, the kernel of the sentence is *John Fox lives*. Students had a hard time understanding this, experimenters had an equally hard time teaching all this all too counter-intuitive stuff⁶.

In Dutch the word *het* is a homonym, being the definite article for the neutrum singular (*the*) as well as the impersonal third person singular (*it*). So, *werkt* indeed is a word that goes with *het* (*het werkt*; Eng.: '*it works*', *it is working*). As the protocol instructed the experimenters to stick to this (obviously idiotic) one-sided word class attack skill (via the article, not via meaning), students as well as experimenters have been known to be at their wits end when confronting this problem (cf. note 6). Trying to form a plural (a cunning trick that

5. In Dutch educational textbooks on grammar the finite verb, more often than not, is considered to be a separate sentence element. One of the things that, in spite of this, makes *Jouw Taal Mijn Taal* such a quality course, is that in this part of the course the predicate in all sentences to be parsed only had one (finite, of course) verb.

6. Small wonder traditional grammar does not enjoy countrywide popularity.

works wonders in Dutch, as it yields, when painstakingly applied, the nonsensical and in English unequalled *de werkten*) usually helped both of them, as a matter of fact.

Nevertheless, neither students nor experimenters felt either positive or negative about this first version of the School Grammar course. Experimenters were content with the protocol. Still, the course underwent some changes. First, some of the material was removed. This was material mainly concerned with morphology or with the spelling of the finite verb (a notorious problem in Dutch). As it had become clear during the pilot study that students had never before heard of some of the subjects that in the School Grammar course were supposed to be familiar (in particular: that sentences could be divided into separate parts that all had a meaning of their own), new material was added to introduce these subjects. Of course, the protocol was adjusted, too. Text and exercises were put together in one book, called *Sentences and parts of sentences* (Du: *Zinnen en delen van zinnen*).

Not surprisingly, the resulting grammar course focuses on parsing sentences and recognizing parts of speech. It does not link parts of speech and sentence elements. Being a typically 'modern' traditional course, it introduces traditional terminology only late or not at all, freely using terms coined on whatever bases that turn out convenient. Most noteworthy of all is the complete avoidance of the word *verb*. *Finite form* (Du: *persoonsvorm*) is the first regular name the students are made familiar with as far as verbs go and that only in the last chapter of the book. Throughout the course verbs are called 'now or yesterday words' or words that tell 'when', thus hammering home the semantics of the finite verb while in the mean time cramming students for one of the crucial heuristics in identifying this verb form. Here are two examples.

(*Sentences and parts of sentences* (SPSS), pages 29 and 30)

Compare sentence 1 and sentence 2.

1. I am in class.

2. I was in class.

Sentence 1 and sentence 2 are not the same completely. One word is different.

* Underline that word in sentence 1 and sentence 2.

* Make other sentences such as 'I am in class' (all now).

3. I am in class.

4. I in class.

5. I walk to the blackboard.

6. I to the blackboard.

7. I

8. You

* Make other sentences such as 'I was in class' (all yesterday).

9. I was in class.

10. I in class.

11. I went to the blackboard.

12. I to the blackboard.

13. I

14. You

(SPSS, page 33)

- * In the now sentences, underline the word that tells that it happens now.
- 1. Willy was an able shoeshine-boy.
- 2. Willy is an able shoeshine-boy.
- 3. He puts down his box.
- 4. He put down his box.

Contrariwise, nouns are introduced by emphasizing formal characteristics solily, viz. as *de*-and *het*-words, *de* and *het* being the two definite articles in Dutch (*het* for neutre, *de* for feminine and masculine). Then again, sentence elements are named after their scope, rendering elements that tell how, elements that tell where, elements that tell what, elements that tell when and so on and so forth. The latter element is, of course, due to clash happily with all verbs dealt with in the course. Treatment of subject material alternates likewise. Here are some more examples.

(SPSS, pages 11, 15 and 16)

The walkers merrily wandered through the quiet wood.

- * Who wandered through the quiet wood merrily? Draw a square around the words that tell who wandered through the quiet wood merrily.
- * Where did the walkers wander? Draw a square around those four words.
- * One word tells how the walkers wandered through the wood. Draw a square around that one word.
- * What did the walkers do in the wood. Draw a square around the word that tells what the walkers did.

A dog sniffed his pants carefully.

- * How did the dog sniff? Fill in the correct answer.

It sniffed

- * Think of three sentences with a howpart.

.....

- * Also think of a sentence with a whenpart. Write the sentence here.

.....

- * Think of a sentence with a whompart. Write the sentence here.

.....

- * Complete these sentences

When?	Who?	What did he do?	Where?
1 In the morning	the engine driver	climbed	to his seat.
2		walked	down the street.
3	puss		in the kitchen.
4 After one hour	father		
5 Last week			at ours.

Material to be dealt with comes in themes, to be covered in three or four classes. Sentences and their elements are, not surprisingly, the central theme. The course has students change word order while retaining meaning of sentences in order to get them to discover what words are in one element, like this:

(SPSS, pages 12 and 13)

Marja sang songs all the way.

- * Make another sentence with these words. The sentence has to start with the word all. Write your sentence here:

.....

* Make another sentence with these words. The sentence has to start with the word *sang*⁷. Write your sentence here:

* Now compare the three sentences. Draw a square around the words that stay together in the three sentences.

In order to get an idea of their individual meaning elements are substituted, added or deleted. Students glimpse at rather more abstract aspects of syntactic structure by generating sentences framed alike. Also, students construct sentences with given elements, moving around small slips of paper individually or instructing Miss to do the same with larger ones on the blackboard. Discussing which are the extra elements, adding detail to the sentence, and which are the ones that cannot be left out, students gradually though resolutely are led towards the notion that some elements are more fundamental than others. They confront the kernel of the sentence as early as the third series of lessons. Here is an exercise that deals with this kernel.

(SPSS, page 101)

1. Rather limply, the hat hang on its peg in the corridor.
 * Shorten this sentence, make it still shorter, and shorter still.
 * Fill in: you are left with the k.....
 * Shorten sentences 2 to 5, make them still shorter, and shorter still. Underline the kernel of the sentence.
2. Sadly the hat was lying amongst the rags on the cart.
 (De hoed lag verdrietig tussen de lompen op de kar.)
3. Mr Drusse appeared on television every Saturday at nine.
 (Meneer Drusse speelde elke zaterdag om negen uur op de televisie.)
4. All clothes were living together in the house on the corner.
 (De kledingsstukken woonden bij elkaar in het huis op de hoek.)
5. The old hat has been lying for years in the wardrobe with all other things.
 (De oude hoed ligt al jaren in de kast bij de andere dingen.)

In preparation to and as a help in recognizing the kernel of the sentence, concord of subject and finite verb is duly stressed. Obviously this is done by having students turn the 'part of the sentence that tells who' from singular into plural (or vice versa), thus calling to attention the word that changes in tandem with it. Here is an example.

(SPSS, page 109)

- Recently I saw a thrilling movie.
- I just could not get enough of it.
- Do you ever listen to a radio play?
- During daytime I hardly ever watch telly.
- * Fill in: Are the subjects of these sentences singular or plural? They are
- * Turn the subjects of the sentences to plural. Write down the whole of the sentence. Be careful! There is another word that changes along with it.

Students familiarize themselves with compound sentences by analyzing these in analogy to their analysis of simple sentences; the latter are made up of elements, the former of sentences. Conjunctions are the words that act as glue.

7. In Dutch this makes a correct question.

Every so often attention is paid to punctuation. Students segment text into sentences by inserting capital letters and full stops or question marks, they segment sentences by putting commas and once or twice they put in colons and quotation marks.

Crucially, all sentences to be parsed or punctuated, all kernels to be isolated or expanded, all frames to be used and all meanings to be explored are not generated by the students but given to them. Equally crucial, explicit rules are stated in rather indefinite and widely applicable terms as students are supposed to discover finer details and exceptions themselves by generalizing spontaneously over all material presented. This goes for parsing as well as for recognizing parts of speech⁸.

A class would usually start with Miss introducing a new topic from the book. Reading a story and discussing it now and then would precede these introductions. A small amount of class work would follow suit. Next, students devoted some twenty to twenty five minutes to seat work. Students might en masse encounter insurmountable problems, resulting in a long queue at Miss' table. In such cases, the problem would be routinely attacked in a joint effort of all students, Miss skilfully guiding the students along all pitfalls.

When no such problems ensued, some of the work would be discussed in some detail. If enough time was left, a new, sometimes related problem would be introduced, discussed and tackled in class. Also, time was devoted to students completing or correcting exercises in their workbook.

4.4 Sentence Construction

The course that comes closest to the kind of sentence construction wished for in the experiment is *Grammatica in Balans* (*Grammar in Balance*) (Van Dort-Slijper et al. 1987). *Grammatica in Balans* is a course in traditional grammar that emphasizes the meaning of words and sentences. Major word classes are attacked via semantics ('with a noun you name a person, an animal or a thing; with a verb you name what they do'). The relation between meaning and structure in the sentence is highlighted by so-called 'sentence patterns'. Such patterns, framed by parts of speech and based upon the sub-categorization of the verb, relate to the meaning of the sentence. Sentences sharing a pattern tell (or ask) similar things. This relation between meaning and structure is kept up in morphology and parsing as well. Moreover, contrary to common practice in traditional school grammar (in The Netherlands), students are trained to construct sentences by using the sentence patterns. Next, they add detail and information to it. So, they do not break sentences down into their constituent parts to be left with this mysterious kernel.

8. One of the experimenters undertook to study whether students' work showed them to do so. It didn't (Plompen 1990).

Grammatica in Balans is a course for the three last years in primary school (groups 6, 7 and 8). In the pilot study the material for the youngest students was used. In theory students can work their way through the course on their own but as this might very well result in enormous and unwished for differences between students in one group, problems were introduced and tackled in class en bloc. Some twenty to twenty five minutes per hour were devoted to this joint activity. Then, students individually started reading things over in their textbooks and making exercises in their workbooks.

It would be typical for a *Grammatica in Balans* teacher to say

- How did we put this sentence together?
- What are the basic words of this sentence?
- Add an adjective to the noun phrase in this sentence to tell what the boy (or anything else that might be mentioned in the sentence) looked like.
- You can tell what things look like by using either a noun phrase (my red socks) or a sentence (my socks are red).

Playful though the course is, with subjects and themes centered around the common or garden hubbub of some four to six children, *Grammatica in Balans* obviously is a grammar course pur sang, relentlessly hammering home grammatical terms and notions, exploiting each and every occasion to meticulously illustrate the tight knot of meaning and structure in language.

Students enjoyed the course enormously and kept on pestering experimenters to stop lingering over things they by now understood very well and to hurry on to the next problem. Experimenters liked the course but did not like the protocol as they felt that that was what made them dilly-dally, causing problems in managing students. So, pace had to be speeded up.

Another, far more serious, problem was that *Grammatica in Balans*, being such a thoroughbred grammar course, concentrated on all kinds of grammatical gadgets of which aspects of sentence construction just was one. The course used did not nearly as often deal with constructing sentences as was desirable. Thus, a completely new sentence construction course was put together by removing all that did not bear on the matter at stake and by adding and elaborating on all that did, even though in the original *Grammatica in Balans* version it was said to be meant for eleven-year-olds. As the newly fabricated course would deal with sentence construction and sentence construction only, including material out of these textbooks was not thought to be all that dangerous. In this way *Grammatica in Balans*, while being reconstructed into *Constructing Sentences*, as the students' book eventually was called (Du: *Zinnen Bouwen*), lost most of its pervasively out and out grammatical character.

It would be typical for a Sentence Construction teacher to say

- (concerning the sentence *Bob is unhappy*) Does this sentence tell what Bob does or how he feels?
- What odd jobs do you do around the house? List two of them.

- These are the notes Bart made during the interview. Turn them into sentences.
- Add extra information to your sentences. For instance, tell what the people or the animals looked like or what you think of them. Use adjectives to do so.

Constructing Sentences, then, though clearly inspired by *Grammatica in Balans*, at last only marginally resembled it.

Text and exercises were put together in the student's workbook. Part of the material was covered in class, experimenter and students working jointly along. Some of the material the students studied on their own.

Students are introduced to studying language by first studying a bicycle⁹. The necessary parts and the extras of both (for bicycles these are, e.g. wheels and handlebars, bell and carrier respectively, in language these are basic and extra words) are dealt with subsequently, students right away constructing sentences by first using basic words only, adding detail in second instance. Here is the exercise.

(*Constructing Sentences (consens)*, page 10)

Jolien wrote six words. She underlined the necessary words.

pink fantastically soap my new lathers

- * How many necessary words do you see?
- * Make a sentence with the necessary words only. That is, a sentence with two words. Mind the capital letter and the full stop.

.....

- * Now tell more in your sentence: add all extra words above.

.....

- * Now make a question with all these words. Mind the capital letter and the question mark.

.....

In order for students to get acquainted with the first of the sentence patterns, nouns and verbs are introduced immediately following this first lesson. Extra words are dealt with in two sessions, one devoted to determiners such as articles, demonstratives, numerals and adjectives, the other devoted to adverbs and adverbials. This is what a typical exercise looks like.

(*consens*, pages 56 and 57)

(At the children's farm)

Fatma and Habibe are watching the kangaroos in the meadow.

1. Kangaroos skip.

N V.

- * Add an extra word to sentence 1a to tell how many kangaroos skip.
- 1a. kangaroos skip.
- * Write the sentence pattern below sentence 1a.
 - * Draw a line around the noun and the extra word, that is around the noun phrase.
 - * Add another extra word. Use the other word to point out what kangaroos you mean exactly.
- Write the sentence down with both extra words.
- 1b. kangaroos skip.
- * Write the sentence pattern below sentence 1b.

9. Dutch as Dutch can be!

- * Draw a line around the noun and the extra words, that is around the noun phrase.
You can tell some more about the kangaroos. For instance, what colour they are or what size they are.
- * Write down the sentence with all extra words.
- 1c.
- * Write the sentence pattern below sentence 1c.
- * Draw a line around the noun and the extra words, that is around the noun phrase.

This series of six lessons lays the foundations for all that is to follow. Most of this expressly reverts to the fundamental notions taught there and then before introducing new subject material.

Students get to know four sentence patterns. Two are patterns that tell what people, animals or things do; two are patterns that tell what they are or what they are like. Here is an exercise introducing one of the patterns.

(*consens*, page 167)

Lottie says to Habibe: 'You are small. Your hair is black. Your eyes are brown.'

- * Now tell what you look like.

I am

My hair is

My teeth are

- * Now also fill in the linking verb.

My eyes

My nose

Framing sentences, students discover these basic patterns remain the same no matter how they expand or change them, changing singular to plural, or present tense to past or perfect tense or declarative sentences into questions.

About one third of the course concentrates on analysis of sentences, asking students to recognize sentence patterns. Such an analysis has its own question-and-answer ritual, the first question always being what it is the sentence tells: does it tell what somebody does or does it tell what he is? The next questions in due order probe deeper and deeper down into the meaning of the sentence, the structure this is framed in and the words that scaffold it. Adverbial elements, on sentence level and thus not included in the basic sentence pattern, are set apart. In sentences that tell what someone is, or is like, the verb complement gets identified first, immediately followed by a search for the word or words that the complement is about (the subject, of course). Only then, students seek for the word linking these words or groups of words. In sentences with a lexical verb, i.e. in sentences that, rather over-generalizing, tell what people do, the predicate is identified first. First focus is on the verb or verb phrase, this being the element that the sentence is framed upon. In case the sentence contains a transitive verb, the object (the noun phrase included in the predicate) is the next element to be identified. The subject is the last element to be taken into consideration. Before running such an analysis on sentences themselves, students take part in an analysis directed by Miss so that the whole of the class, as it were, gears into the job. When all students are able to recognize the sentence pattern at first glance, the ritualized questioning-and-answering is abbreviated by the by. Here is an example of an exercise that tries to carefully guide students in recognizing sentence patterns.

(COURSE, pages 200-202)

At last the day has arrived: it is Joy's birthday.

What do the children do on Joy's birthday? First they all go to the fair: Joy and Jolien and Isabelle and Lotte and Klaas and Gijs and Oğuz and Fatma and Lennart. And Joep? Joep does not¹⁰.

1. Lotte, Oğuz and Jolien are loitering about the fair.

(Lotte, Oğuz and Jolien slenteren over het kermisterrein.

- How many children sentence 1 is about? Fill in: sentence 1 is about children.
- What do we tell in sentence 1? Fill in: we tell what Lotte, Oğuz and Joy
- In sentence 1, do we tell where they are loitering? Tick the correct answer.
 - 0 Yes.
 - 0 No.
- In sentence 1 draw a square around the words that tell where they are loitering, that is, around the extra phrase.
- Write the correct sentence pattern below sentence 1.
- 2. The children are teasing Nelson and Timo at the shooting gallery.
- What do we tell in sentence 2? Fill in: we tell what the children
- In sentence 2, do we tell where they are doing this? Tick the correct answer.
 - 0 Yes.
 - 0 No.
- In sentence 2, draw a square around the words that tell where they are teasing Nelson and Timo. That is, around the extra phrase.
- Write the correct sentence pattern below sentence 2.
- The children head for the roller coaster. Jolien and Lennart find two empty seats.
- 3. Jolien and Lennart are cheerful at first.
- 4. Lennart suddenly becomes afraid and sick.
- 5. Is the roller coaster nice or frightening?
- What do we tell in sentences 3 and 4? Tick the correct answer. We tell
 - 0 what Jolien and Lennart do.
 - 0 what Jolien and Lennart feel like.
- What do we ask in sentence 5? Tick the correct answer. We ask
 - 0 what Lennart and Jolien think of the roller coaster.
 - 0 what the roller coaster is doing to them.
- In sentences 3 and 4, do we tell when Lennart and Jolien feel like this? Tick the correct answer.
 - 0 Yes.
 - 0 No.
- In sentences 3 and 4 draw a square around the words that tell when they feel cheerful, afraid and sick. That is, around the extra phrase in sentence 3 and around the adverb in sentence 4.
- Write the correct sentence patterns below sentences 3, 4 and 5.

One third of the course has students create or elaborate sentences themselves. As a rule, they are given general instructions as to how go about this. In describing one or two elements of a drawing, e.g., they have to pattern a basic sentence on one of the sentence frames and step by step add extra information to it, such as number, colour, material, quality or precise whereabouts. For a finger exercise, the whole procedure would be gone through collectively in class, students supplying alternatives only, Miss making an inventory of some of the things they come up with on the blackboard. Notice that as this activity concerns single, isolated sentences only, students are never engaged in actual writing (cf., however, chapter 2.6 on the supposedly a-rhetorical context of courses such as these). As framing sentences in this way does not always yield sentences open to straightforward student analysis, no retrospect analysis is run on these. The following are examples of such sentence construction activities.

10. As Joep is a dog, he wouldn't like the fair anyway, would he?

(consens, pages 157 and 158)

Jolien is staying on the farm with her cousins Jelle and Homme. Homme is four years old. He is going to primary school.

'Do you like going to school?' Jolien asks.

'Sure', Homme says, 'I want to learn how to read.'

'Shall I read you something?' Jolien offers.

'I only have three books and I know them by heart', Homme sighs.

'Then I will think of something myself,' Jolien says. 'The sentences have to be simple, so you will be able to read them by yourself in a while.'

Jolien writes a verse:

Cocks crow.

Cows low.

Muck spreaders blow.

The land is being manured.

The watchdog barks

while Homme sleeps.

Jolien thinks the sentences are short indeed. She likes to tell more about the animals and things in her verse. For example, what the animals look like, or how many there are. So she adds two extra words to all nouns. You help her.

..... cocks crow.

..... cows low.

..... muck spreaders blow.

The land is being manured.

The watchdog barks

while Homme sleeps.

(consens, pages 160-162)

(Did you have a good look at the drawing on page 159?)

• Write down two nouns to name what you see in the drawing.

1.

2.

• Add an extra word to each noun. For example *de* or *het* (Eng: *the*) or *this* or *two*. You may think of an extra word yourself. Write the extra word and the noun in 1a and 2a. There are two noun phrases, now.

1a.

2a.

• Now tell what the people, animals or things look like. Add an adjective. The noun phrases grow.

1b.

extra word adjective noun

2b.

extra word adjective noun

• Now tell what you think of the people, animals or things. Add another adjective. The noun phrases get bigger still.

1c.

extra word adjective adjective noun

2c.

extra word adjective adjective noun

• Now make sentences. Tell what the people, animals or things do. Take care your sentences look like sentences.

1d.

extra word adjective adjective noun verb

2d.

• Can you write the sentence pattern below sentence 2d?

• In sentences 1d and 2d draw a line around the noun phrase.

(consens, pages 282-284)

Oguz is writing something for the school newspaper. He writes about Ruud Gullit, the football payer. Here is his first sentence.

1. Gullit is a footballer.

• What is it Oguz tells in sentence 1? Tick the correct answer. He tells

0 what Gullit does.

0 what Gullit is.

• Write the correct sentence pattern below sentence 1.

Oguz is not yet quite happy about sentence 1.

'Gullit is not just any footballer', he thinks. 'I am going to tell what I think of Gullit and what he looks like.

• In sentence 1a, tell what you think of Gullit and what he looks like. Add two adjectives to the sentence.

1a. Gullit is a footballer.

In the second sentence of his writing Oguz tells what Gullit does not do.

2. He does not comb his hair.

(Hij kamt zijn haren niet. Lit: He combs his hair not.)

• Draw a square around the adverb that makes the sentence negative.

• Write the correct sentence pattern below sentence 2.

Oguz is not quite happy about sentence 2 either. He likes to tell how long is Gullit's hair and what it looks like.

• In sentence 2a tell how long is Gullit's hair and what it looks like. Add two adjectives to the sentence.

2a. He does not comb his hair.

In the third sentence Oguz asks what Gullit does do, perhaps.

3. Does such a footballer exercise?

(Traint zo'n voetballer? Lit: Exercises such a footballer?)

• Write the correct sentence pattern below sentence 3.

Oguz wants to be more precise in his question. He wants to know when or how often such a footballer exercises.

• In sentence 3a, ask when or how often such a footballer exercises. Add an adverb or an extra phrase to the sentence.

3a. Does such a footballer exercise ?

Finally, Oguz states in one sentence what he thinks of Gullit. He broods a long time. What exactly does he think of Gullit? Is he super? Wonderful? Nice? A hero? An example? A smashing guy?

• What do you think of Gullit? Write it down in sentence 4.

4. Gullit is

• What sentence pattern comes below your sentence 4? Tick the correct answer.

0 N IV A.

0 N IV N.

Finally, one third of the course deals with the formal characteristics of fully-fledged sentences, highlighting the differences between noun phrases and full clauses. Here is an example.

(*consens*, page 325-327)

Jolien and Lottie have been to the clothes' museum, for they are working on a project on clothes.

Jolien visited the section on clothes of old. She took notes of what she saw. Here are some of her notes.

stuffy, white, powdered wigs
odd, green, mid-length, breeches
long embroidered pink dresses
ugly high black patent leathers
big red adorned hats

Jolien wrote down the colour of each thing.

• In the notes of Jolien, underline the five words that name what colour the things are, that is, underline five of the adjectives.

• Jolien's notes are no complete sentences yet. What are they, then? Fill in: they are

Now Jolien is busy writing. She is turning her notes into complete sentences. In all sentences she tells what colour were the things she saw.

• You help Jolien. Turn the noun phrases into complete sentences. Tell what colour the things were. Take care the sentences look like complete sentences.

Those stuffy powdered wigs were white.

.....
.....
.....
.....

Lotte visited the section on modern clothes. She took notes, too, on all there was to see. Here are some of her notes.

stained cotton jeans
expensive modern sports shoes
tight leather trousers
baggy printed shirts
shiny flimsy miniskirts

'You don't have to visit a museum to see this,' Lotte thinks. 'You can see it on the street as well.'

Lotte gets started, too. She adds something to the information on her piece of scrap paper: she tells what she thinks about the clothes.

* You do that, too. In a sentence, tell what you think of the clothes Lotte saw in the museum. Write down the sentences below. In each sentence there should be three adjectives.

Such stained cotton jeans are beautiful.

.....
.....
.....
.....

As a matter of course, sentence patterns are crucial here. Students turn noun phrase into full clauses and evaluate their formal differences in terms of (the absence of) a sentence frame or, leaving all formal characteristics aside, they are asked to turn notes, taken by one of the protagonists of the course book, into sentences. When directed in their generating a sentence themselves they might also be asked to take down one or two of the things their sentence will be about first and next construct a sentence using this noun phrase.

Punctuation is checked systematically for each sentence or sets of sentences that students write, emphasizing that a complete sentence is always recognizable in writing.

The general set up of Sentence Construction lessons resembled that of the other grammar course, classes starting with Miss introducing a new topic from the book, followed by some ten to fifteen minutes of class work, the routine for en bloc student stumbling being the same as in the course on school grammar. Likewise, when no such stumbling occurred, some of the work would be discussed in some detail at the end of the class or a new, related problem would be introduced, discussed and tackled. Time was devoted to students completing or correcting exercises in their workbook, too, exactly as in the School Grammar course.

4.5 The experimenters

The experimenters-to-be all three were females in their twenties. Two were bachelors of educational linguistics, one was a trained teacher and a student of pedagogy. All had ever in one way or other been working with children of nine years old or thereabouts. They participated in the pilot study as trainees. A more or less official evaluation of their doings was performed by talking over weak and strong points in their performance in class, inevitably stressing the points that had to be changed for the better.

As eventually the three of them turned out to be experts on the job, they were appointed on the research staff as experimenters for the year the actual experiment ran.

4.6 Conclusion

This chapter reported on construction, try out and revision of the three experimental programmes and the protocols accompanying these. On the basis of the information the pilot study yielded all programmes were changed, resulting in three courses that are to be considered exponents of different views on teaching specialties of language arts. All three experimental courses have one thing in common: they all mean to beneficially affect student writing. However, they go about it very differently. And so their results will differ as well.

Most importantly, one course has students write, two have them study grammar. Direct Writing (course book: *Workbook Writing*) trains students in producing texts in diverse genres. The course, however, withholds essential information on the formal characteristics of written language. Dialogues are dealt with, and so are the inverted commas that go with it, but sentences are not. Students discuss start, body and end of a text, they familiarize themselves with paragraphs, they know what is in a paragraph but they do not know its basis unit: the sentence. Students write and the course teaches them how to go about it. It has them draft and revise their prose. It loosens brainstorming to gear them and provides ideas to chew on, comment to think over and spurs to elaborate. This due attention to the contents of their writing will eventually enhance its overall quality.

Both grammar courses come with a load of grammatical cant. Most overlaps, some of it is decidedly idiosyncratic. Both, however, provide students with a solid body of academic linguistic knowledge. As far as that is concerned, then, there are no differences between these two courses. In the School Grammar course (course book: *Sentences and parts of sentences*) students are taught to parse sentences and to find the kernel of the sentence. They are taught how to recognize parts of speech as well. As, however, the course does not touch upon anything having to do with written language, students are hypothesized to come to nothing when sentence awareness or conciseness are concerned. Quality of their prose will be affected neither. The course in Sentence Construction (course book: *Constructing Sentences*) not surprisingly highlights constructing sentences. The course concentrates on the inter-relatedness of meaning and structure by familiarizing students with sentence patterns. So, though the academic linguistic knowledge of Sentence Construction students equals that of School Grammar students (and, of course, outstrips that of Direct Writing students), yet Sentence Construction students are the only ones who know what distinguishes sentences from non-sentences. Therefore they will use more genuine sentences in their writing than do School Grammar and Direct Writing students.

Also, Sentence Construction students will be better able to tightly pack noun phrases than School Grammar and Direct Writing students and they will employ complexly structured noun phrases in their writing more frequently than will School Grammar and Direct Writing students. Their expert knowledge on sentences, culminating in their rightly exploiting these to scaffold their writing will cause Sentence Construction essays to be of an overall quality superior to those of Direct Writing students.

FIVE DEPENDENT VARIABLES MEASURE FOR MEASURE

'... but a Verb is a word as signifies to be, to do or to suffer (which is all the grammar, and enough too, as ever I was taught); and if there's a Verb alive, I'm it. For I'm always a-bein', sometimes a-doin', and continually a-sufferin'.'

Mr Tapley in Charles Dickens, Martin Chuzzlewit.

In the experiment carried out to study the effects of grammar instruction upon junior writing, five hypotheses are put to the test (cf. chapter 3.2. and 4.6). Each hypothesis concerns one of five dependent variables. These are, in Indian file:

- writing fluency;
- academic linguistic knowledge;
- sentence awareness;
- conciseness;
- overall writing quality.

This chapter reports on the construction, try out and revision of the measuring instruments for the dependent variables. First, general information on the collection of data in the pilot study will be given in section 1. In the sections 2 to 6 information on all measuring instruments will be given per variable. Section 7 summarizes this chapter.

5.1 Pilot study

In order to test the hypotheses on the effects of grammar instruction upon writing of primary school students, both knowledge of students and the use they put it to had to be measured. Instruments to measure the knowledge present all came in the form of discrete point tests. Whether this knowledge was available and accessible was gathered from actual student writing. All instruments were tried out in advance. There were pilot studies for discrete point tests and writing assignments and for rating syntactic and overall quality of writing. All will be discussed in this section. For a more detailed report readers should consult Van de Gein (1989).

Discrete point tests

Discrete point tests have been tried out twice. In the first pilot study, data on a total of 14 trial tests, totalling 1123 test items, were collected; items had to be tested for nine- and ten-year-olds, i.e. the youngest and oldest students to be enlisted in the experiment. For estimates of item characteristics to be reliable, some 200 students had to work on each item. All this meant that some 80 primary schools with about 25 students per class were involved in the pilot study, subjects totalling about 2000 nine-year-olds and 2000 ten-year-olds.

Nine continuous working test assistants, equipped with test protocols and trunks, collected all data within two months. Subsequently a team of coders, working from detailed protocols, prepared results for data analysis. Data analysis yielded information for both groups separately on reliability per test (coefficient alpha), and on difficulty or item mean (p-value) and discrimination (rit-value) for each item.

On basis of these results large sets of items within tests had to be removed from the test battery. A total of 695 items remained and were tried out again. In the mean time, students had aged, of course, so only the nine-year-olds (by then being nine-and-a-half-year-olds) were involved in this pilot study. Collection, coding and analysis of data ran the same course it did in the first pilot study. Tests to be used in the experiment were selected on basis of difficulty or item mean (in order to avoid ceiling effects, p-value maximum was set at .75) and on reliability (in order for the ultimate ten item tests to be psychometrically reliable, estimated coefficient alpha for these had to amount to .70 at least). Four tests survived: one on academic linguistic knowledge, one on conciseness and two on sentence awareness.

Writing assignments

Suitable writing assignments were selected by first collecting over 600 essays on each of fifteen topics. There were four trios of topics belonging to four different genres. Three topics left students free to choose genre. For convenience' sake the latter is called the free genre. Genres (and topics) were:

- a call (for a jumble sale, a book fair or a clothes fair);
- a description (of mugs, rag dolls or children's portraits);
- a narrative (on rain boots that went for a walk, a locomotive that went its own way or a whistling kettle that wished to join the school's orchestra);
- an exposition (on favourite hobbies, pets, or either books or programmes on radio or television);
- free genre (a story on something sad, something silly or something thrilling).

Essays were written by all students involved in the first pilot study. Test assistants left assignments with class teachers and these made each student write three impromptu essays: one essay in one genre on three consecutive days. A random selection of ten to fifteen percent of these was analyzed to evaluate the amount of text per genre and comparability of texts within genre and, within genre, of topics. As it turned out, if the free genre and the call yielded texts at all, these were pretty heterogeneous and diverging to such degrees on all

imaginable dimensions as to render them completely incomparable even within topics. The other three genres (Description, Narrative and Exposition) yielded comparable texts, but not all topics turned out to be fortunate. So, topics had to be reformulated and tried out again. Moreover, for each genre there had to be five topics available (one for each test period; cf. chapter 3.3). All in all, in the second pilot study three revised and five new topics were tried out for the three successful genres by having all students write an impromptu essays again. This yielded 200 essays per topic. Revised topics were:

- description of a mug;
- exposition on a favourite book and on a favourite programme on radio or television.

New topics were:

- description of ice creams and of chairs;
- narrative on the baby ghost that was afraid of the dark and on the shop-window mannequin that wanted to celebrate her birthday;
- exposition on a best friend.

A random set of ten percent of all essays was again analyzed on amount and comparability of texts. Both turned out to be satisfactory.

Here is one example for each genre.

Description

Have a close look at the ice creams on the next page. Choose one. Tick it. Then write *as precisely as possible* what the ice cream looks like. Don't forget anything, mind you!

(Bekijk de tekening van de ijsjes op de bladzijde hierachter heel goed. Kies dan één ijsje uit. Zet daar een kruisje bij. Schrijf dan zo *precies mogelijk* op hoe dat ijsje eruit ziet. Niks vergeten, hoor!)



Narrative

Write a thrilling or silly story on the shop-window mannequin that wanted to celebrate her birthday.
(Schrijf een spannend of gek verhaal over de etalagepop die haar verjaardag wilde vieren.)

Exposition

Write something about your best friend. Write, for example, what he or she looks like and why you think he or she is such a fine friend.

(Schrijf iets over je beste vriendje of vriendinnetje. Schrijf bijvoorbeeld hoe hij of zij eruit ziet en waarom je hem of haar zo'n fijne vriend of vriendin vindt.)

There also was a guided test to assess student writing, pilot tested twice along with all the other tests and screened in the same manner the writing assignments were screened. This test made students describe, one sentence at the time, separate elements of people, animals, plants and buildings, drawn in the test booklet. There are five booklets with fourteen items each. Below are two random examples of items in this test.

Here you see a drawing by Jùliette. Take a close look, for there is a lot to see. Now describe what you see. There is one example sentence. (*Three fat white chickens are scratching about in the grass.*)

Now write one sentence on the helicopter.

(Hiernaast zie je een tekening die Jùliette gemaakt heeft. Bekijk de tekening goed, want er is veel te zien. Beschrijf jij nu wat er te zien is. Eén zin is al voorgedaan. (*Er scharrelen drie dikke witte kippen in het gras.*))

Schrijf jij nu één zin over de helicopter.)



Here are four drawings. Each is a drawing of a plant. All plants are different. Compare, for instance, their stalks. Or their leaves. Or their flowers. Now write about the differences you see. First write about the stalks of the plants. There is an example sentence on the plant in drawing 1. Now you write one sentence on the stalk of plant 3.

(Hier zie je vier tekeningen. Het zijn tekeningen van planten. Elke plant is anders. Vergelijk de stelen maar eens. Of de blaadjes. Of de bloemen. Schrijf nu iets over de verschillen die je ziet. Schrijf eerst iets over de stelen van de planten. Er is één voorbeeldzin over de plant in tekening 1. Schrijf jij nu een zin over de steel van plant 3.)



Administering of assignments and analyses of texts

In order to control for topic effects, all 70 items of the guided writing test were allocated randomly to test booklets. Result was that students encountered all drawings more than once (there were seven in total) but had to describe different things each time.

Topic effects of descriptions, narratives and expositions were controlled by allocating topics to students and test periods systematically. That is to say: all topics were written on simultaneously but not, of course, by all students. At the time of the pre test, for instance, one fifth of the Sentence Construction group

at school 1 (i.e.: one or two students) and one fifth at school 2 (two students) and at school 3 (one student) and at school 4 (two or three students) and at school 5 (one student) and at school 6 (two or three students) wrote a Description on ice-creams. So did one fifth of all Direct Writing and School Grammar students at schools 1 to 6. The second fifth of all experimental groups per school wrote another description, on, e.g. rag dolls, the third fifth still another and so on. Narrative and expository topics were allocated likewise¹.

Measures for analyzing syntactic aspects of student writing were tried out on a small scale by having five professional linguists work with a protocol and score these aspects on a corpus of 60 essays, written in the pilot studies. Inter-rater reliability was explored and found to be ok. The protocol was revised extensively on basis of detailed analysis of results. This revised version was not tried out again. Details are presented in sections 2, 4 and 5 of this chapter.

In this research the syntactic aspects of writing are highlighted. Non-syntactic aspects of writing are only glimpsed at and this is done in a rather global way, too, viz. by considering the overall quality of an essay. There is as yet no unanimity among writing researchers on what exactly constitutes overall writing quality, or, in other words, what gets rated when essays are rated holistically. Section 6 of this chapter will touch upon some of the issues at stake.

Scoring of essay quality was tried out only in order to establish whether raters could do without a protocol. They could. For details, also consult section 6 of this chapter.

Though genre effects will be explored per measure, measures will be computed over all non-guided writing on the one hand and the guided writing test on the other. No evidence whatsoever was found to imply that such a summation over these genres is untoward.

5.2 Writing fluency

Writers draw on their tacit linguistic knowledge. So, says the first of the set of hypotheses, writing fluency is fine and no treatment can, will or need affect it. The instrument to measure writing fluency rests on the so-called T-unit (Hunt 1965; a T-unit is, roughly, a sentence including all subordinate sentences). Traditionally, in an analysis in terms of T-units at least mean T-unit length and number of clauses per T-unit are computed. Both measures have come to be known as reliable indices for syntactic growth (whatever that may be, also see chapter 2.6). To measure T-units, one ignores the punctuation (if any) in the

1. I am not exaggerating when I say that this was a hell of a job as, in practice, it meant putting to name all essays to be written. There were 15 essays to be written and over 165 names. Mr. Gerrit Jan van der Duim supervised this circus and thanks to his extreme scrupulousness nothing, but nothing went wrong. All honour to him!

essay under investigation. Instead one identifies all syntactic strings according to certain instructions, of which the aforementioned one is central. Generally, in a T-unit analysis, if it is easy to see (or hear) what is wrong in a deviant string, this string gets repaired and the repaired string is included in the analysis. Strings beyond repair (known as 'garbles' or 'mazes') are not included in analysis. As measuring students' performance this way is actually said to give insight in linguistic competence of language users, this repairing or deleting of material allegedly is the natural thing to do as, in that case, one is wont to explain away performance errors.

Students enrolled in the study into the effect of grammar instruction upon their writing met with no such benevolence. Rather, all they wrote could be (and indeed was) used against them. There is one enormous drawback to such rigorous stringency: deviant strings are all tarred with the same brush. Strings lacking an in all probability in haste forgotten element; strings very likely defective either due to students' eagerness to write or due to their sheer inability to keep themselves going; aberrant strings, looking as though they were the products of some hard conscious monitoring; all are tried and found wanting. Accidental errors resulting from a momentary attention dip and non-accidental errors, resulting from sheer lack of knowledge are indiscriminately thrown on the same heap of flawed strings.

Though lumping all this together under the heading of faulty strings may well seem harsh and inaccurate, it is precisely what should be done. Writing fluency of novices, says the hypothesis, is fine, which amounts to saying that errors are marginal and need not be bothered with. By setting apart flawed strings it will be possible to show how fine it is. The hypothesis also states that writing fluency will nor can or need be affected by treatment. Separating flawed and non-flawed strings will also bring out whether there is any truth to this hypothesis.

To establish writing fluency of students, all T-units were analyzed in the exact form that they were written down in. Subsequently, the amount of text that came in flawless strings was calculated. This was done by computing the T-unit ratio, i.e. the ratio of the total number of words in flawless strings and the total number of words in the essay².

In the small scale try out that was run on this measure (see section 1 of this chapter), scrupulous analysis of data revealed that raters, albeit all were professional linguists, seriously disagreed as to what was or was not a T-unit (i.e., what was or was not an independent sentence and what were the sentences dependent on it)³. In order to minimize misconceptions about this, all essays

2. Theoretically, words are not the most charming elements to work with in such an analysis. After all, they do not seem to be what writing is either produced or received in. For practical purposes, however, they are unequalled.

3. So even to these most ideal of all ideal speakers/hearers, sentences do not come as naturally as all that.

were prepared for analysis by dividing them up into T-units⁴ (see appendix 1 for the rules used for this).

Raters' agreement on what was to be considered a flawless string and what wasn't was satisfactory. Still, a revised version of the protocol provided raters with rather extensive rules for this. These rules stated, basically, that flawless strings had to have all necessary words in it and in the correct order as well. Incorrect spelling, incorrect word formation (such as the actually encountered *dronkenheid*, Eng: *drunkenship*), incorrect conjugation or declination (of verbs and adjectives respectively) and incorrect gender (in Dutch manifest by the presence of an incorrect article) did not matter and had to be neglected. Raters were warned that as all strings were analyzed out of their context some strings might seem to be incomplete whereas they actually were not so. Answers, direct questions in a dialogue (with or without inverted commas), greetings, stereotyped expressions (such as *no sooner said than done*, Du: *zo gezegd, zo gedaan*), explanations (such as one student gave: *george, spreek uit djoorts*, Eng: *george, pronounce djoorts*, which, in Dutch, is correct!), exclamations, sighs or sounds (*toot toot toot*) all had to be considered correct. Also, raters were warned that erroneous ordering of adjectives within noun phrases had to be neglected and that marked structures such as in narratives (e.g.: *In comes an enormous giant!*, Du: *Komt me daar een grote reus binnen!*) and jokes (e.g. *Says one to the other*, Du: *Zegt de een tegen de ander*) are correct.

Flawed strings, then, are incomplete or incorrectly structured ones, strings with a blending error, strings that are typically oral language, and, of course, all strings harbouring several of these flaws simultaneously.

In order to paint a complete and lively picture of writing fluency of juniors three measures were added. One gathers information on the mean length, or, indirectly, on complexity of flawless strings. The other two bear on rather stylish aspects of writing fluency. One, the run-on ratio, concerns the way writers conjoin T-units. Young writers are renowned for their tendency to 'run on' in essays. By using numerous conjunctions they coordinate numerous strings, in some essays culminating in one bulky 'sentence'. Run-on sentences, as these have come to be known, are common in and, in all probability, characteristic for casual speech and thus, too, common in and characteristic for junior prose. The run-on ratio is the ratio of all strings starting with *and*, *but* and *then* and the total number of strings.

The other stylish aspect of writing fluency concerns the noun phrase, pronouns excepted. One measure explores the popularity of the noun phrase with novice writers as a substitute for full clauses. The ratio of noun phrases embedded in a flawless syntactic string and those not so embedded was the measure to glean insight into this aspect of writing fluency.

4. By yet another ideal of ideal speakers/hearers, viz. by me.

To conclude, the quartet that measure writing fluency are

- T-unit ratio;
- mean T-unit length;
- run-on ratio;
- ratio of noun phrases per T-unit.

All measure aspects of syntactic strings in writing. As the latter three are merely meant to explore minor details, the first one is the only one that will be used to test the hypothesis that writing fluency of students is fine and that no treatment will affect it.

As an illustration to what these measures assess, the four of them are set at two essays. The translated versions of both are presented below. Numbers mark T-units; indented T-units are part of one studentwise demarcated sentence that stretches over more than one T-unit; noun phrases come in square brackets. Scores on all four measures for each essay immediately follow it.

(5.1.) Narrative on the shop-window mannequin that wanted to celebrate her birthday⁵

- 00 (Once upon a time there was a doll in a shop-window)
 01 [Night and Day] she stood in [the window].
 02 [Nobody] bought [the doll].
 03 She had been standing there for [half a year].
 04 [Such a lot of children] had laid [their hands] on her.
 05 But there was always [something] amiss with [the doll].
 06 Either it was too expensive or she was dirty.
 07 [One day] it was [the doll's birthday].
 08 But [no-one] said to her: Congratulations.
 09 There she was, [the whole of the day], in [the window].
 10 She wanted to celebrate [her birthday] so very badly,
 11 but she knew that [such a party] was out of [the question].
 12 When it was dark she went to have a look in [the shop] all the same.
 13 And at last [the teddy bear] came to her and said: Congratulations.
 14 She walked down [the shop] with [the teddy bear].
 15 She got [a new dress] [new ribbons for [her hair]].
 16 She looked very beautiful.
 17 There was [a very big and beautiful party].
 18 She had [a piece of [birthday cake]].
 19 They were dancing and singing.
 20 Until it got light again.
 21 Then [the doll] had to go back and stand in [the window].
 22 And then came [the biggest present of [her whole life]].
 23 [A girl] came walking towards [the doll] and said to [her mother]: "[Mummy], this is [the doll] I
 want to have for [my birthday].
 24 Yes, it certainly is [a beautiful doll].
 25 [The salesman] said: How can this be?
 26 [Yesterday] she wore [another dress].
 27 How come I didn't put [these clothes] on her before?
 28 I certainly would have sold it earlier.
 29 You're lucky, [madam].
 30 [yesterday] I reduced [this little doll] in [price].
 31 You can [this doll] for [20 guilders].
 32 Yes, I'll have [the doll].

5. Dutch for *shop-window mannequin* is *etalagepop*. It turned out students took *etalagepop* to be a homonym, meaning either a shop-window mannequin or a doll displayed in a shop-window. Students' interpretation of the word was no problem in writing or in analysis.

33 And so [the doll] was sold.
 34 Furively [the doll] waved to [the teddy bear and the other toys] and took [her leave].
 35 Then [the girl] walked out of [the door of the shop].
 36 And [the next day].
 37 Was [the girl's birthday].
 38 She was given [the little doll] by [her mother and father].
 39 She was very happy with it.
 40 And she put [the doll] on [her bed].
 41 And [the little doll] was very very happy,
 42 and so was [the girl].
 (4/31/33/M5; Petra)
 T-unit ratio: .98
 mean T-unit length: 8.45
 run on-ratio: .29
 ratio of noun phrases per T-unit: .97

(5.2) Exposition on a favourite television programme

00 [handling horse and pony]
 01 This is [a programme with [all kinds of horses and ponies and Shetlanders and lots of [other
 horses]]]
 02 and how you should handle [horses]
 03 and how you should ride [a horse or a pony]
 04 and that you should not be behind [a horse or pony]
 05 for if it shies it may kick backward
 06 and they also show [all horses and ponies] that are there
 07 and what you should not do
 08 and how you should put on [a halter]
 09 and [lots more] you need to know
 (2/12/14/M5; Debora)
 T-unit ratio: 1.00
 mean T-unit length: 9.44
 run on-ratio: .78
 ratio of noun-phrases per T-unit: 1.00

5.3 Academic linguistic knowledge

One of the hypotheses of this study holds that academic linguistic knowledge of Sentence Construction students equals that of School Grammar students and out-strips that of Direct Writing students. To test this hypothesis, tests on recognizing parts of speech had to be framed. Three multiple choice tests using two different question formats were made for recognizing parts of speech: one for nouns, one for verbs and one for adjectives. As no grammatical terms could be used, all tests were introduced by test assistants, using detailed protocols. Reliability turned out to be dependent on question format. In the most reliable format (estimated mean alpha for 10 items: .62) students had to tick one of four words.

This format was used to reconstruct all three tests. The second pilot study brought out the overall item mean of the three tests was unchanged ($p = .44$) and reliability had increased only for the verb test (estimated alpha for 10 items: .69). Though this reliability stayed below minimum level (set at .70) it was still decided to include this test, Recognizing Verbs, in the test battery.

Five parallel test versions were created by randomly assigning items equal in variance, p- and rit-value to versions I to V. To guarantee reliability each separate version of the test was lengthened by four items randomly chosen from the test version preceding it. For the version to be used in the first test period four new items were made up. The four extra items together came last in the test materials on a separate sheet so as not to soil the proper test items by memory effects.

Here are three random items from the Recognizing Verbs test.

Below are four words in a row. Three words are related. One word is different. It should not be there. Tick that one word.

(In het rijtje hieronder staan vier woorden. Drie woorden horen bij elkaar. Eén woord is anders. Dat hoort er niet bij. Zet een kruisje voor dat ene woord.)

- a. saw (v)
- b. walked
- c. looked
- d. supid

- a. zag
- b. liep
- c. keek
- d. stom

- a. asks
- b. says
- c. wrong (a)
- d. rectify

- a. vraagt
- b. zegt
- c. verkeerd
- d. verbeteren

- a. mast
- b. washes
- c. splashes
- d. scrapes

- a. mast
- b. wast
- c. plast
- d. kraast

5.4 Sentence awareness

Two discrete point tests were framed to measure whether students knew what distinguishes a fully-fledged sentence in written language from strings that in spoken language sail under sentence colours. These mean to test the hypothesis that Sentence Construction students know what distinguishes sentences from non-sentences better than School Grammar or Direct Writing students do. One of these is the Recognizing Sentences test. It is a multiple choice test that has students either tick the one correct sentence (out of five alternatives) or the shortest but still complete sentence (out of four alternatives) or the two or three words (out of five alternatives) that have to be included in a sentence necessarily. This test, containing 79 items, turned out to be fine rightaway. Difficulty or mean p-value was .53, estimated reliability for ten items was .76. No further try-out was done on this test. Fifty items with highest rit-values were selected to be included in the definite test. Five parallel test versions were made by randomly assigning items equal in variance, p- and rit-value to versions I to V. Here is an example for each of the question formats.

Below you see groups of words. Once these are a sentence. Tick that one sentence.

(Hieronder zie je woorden bij elkaar staan. Eén keer vormen die woorden een zin. Zet voor die ene zin een kruisje.)

- | | |
|------------------------------|---------------------------|
| a. Tanja is making pan-cakes | a. Tanja bakt pannekoeken |
| b. Beating the batter | b. Het beslag kloppen |
| c. Melting the butter | c. De boter smelten |
| d. Flour and eggs | d. Meel en eieren |
| e. Big eats | e. Smikkelen en smullen |

Below tick the sentence that is shortest but still complete. Tick the shortest complete sentence.

(Kruis hieronder de zin aan die het kortst is, maar die nog wel volledig is. Kruis de kortste volledige zin aan.)

- | | |
|------------------------------|---------------------------------|
| a. Huub | a. Huub |
| b. Huub laughed | b. Huub lachte |
| c. Huub didn't laugh | c. Huub lachte niet |
| d. Huub laughed all day | d. Huub lachte de hele dag |
| e. Huub didn't laugh all day | e. Huub lachte de hele dag niet |

Below you see groups of words. What words do you need to make a complete sentence? Only tick two letters. Do not make up new words yourself.

(Hieronder zie je woorden bij elkaar staan. Welke woorden heb je beslist nodig om een volledige zin te maken? Je mag maar twee letters aankruisen. Je mag er zelf geen woorden bij verzinnen.)

- | | |
|-------------------------|--------------------------|
| a. that big red balloon | a. die dikke rode ballon |
| b. over our house | b. boven ons huis |
| c. through the clouds | c. in de wolken |
| d. is floating | d. zweeft |
| e. for an hour | e. al een uur |

Below you see a sentence. Make it as short as possible. You need to be left with a sentence, though.

Tick the letters below the words that together make such a short sentence.

(Hieronder zie je een zin. Die zin moet je zo kort mogelijk maken. Je moet nog wel een zin overhouden.)

Kruis de letters aan onder de woorden die samen zo'n korte zin maken.)

- | | | | | |
|-----------|-------|-------|-------------------|------------------|
| In spring | loads | crawl | through the grass | with one another |
| a | b | c | d | e |
-
- | | | | | |
|-----------------|---------|--------|------------|---------------|
| In het voorjaar | kruipen | padden | met elkaar | door het gras |
| a | b | c | d | e |

The second is a test on Constructing Sentences. This is an open ended test that asks students to construct complete sentences with two or three out of minimally five given constituents. The questions came in two formats. The first pilot study yielded no reliable results on this test: hardly any student had worked on it. In the second pilot study, however, the test, containing 50 items selected after scrutinizing the very few data, turned out to be fine (p-value: .66; estimated alpha for ten items: .74). Five parallel versions of this test were made in the aforementioned manner. Here are two examples of each question format.

Turn the words below into a statement or into a question. Your sentence should be as short as possible. So only use the words you absolutely need. Do not change the words.

(Maak met de woorden hieronder een vertelzin of een vraagzin. De zin moet zo kort mogelijk zijn. Gebruik dus alleen de woorden die je beslist nodig hebt. Je mag de woorden niet veranderen.)

the angels - these children - in the play - will be
(de engelen - deze kinderen - in het toneelstuk - worden)

today - these trees - will - all - be chopped down
(vandaag - worden - die bomen - allemaal - omgehakt)

Below you see cards with words. Choose three* cards that make a statement or a question. What sentence do you make with those two cards? Write down that sentence.
(Hieronder zie je kaartjes met woorden. Kies er drie* kaartjes uit die een vertelzin of een vraagzin vormen. Welke zin kun je maken met die kaartjes? Schrijf die zin op.)

stroke the small red pussy over there that big dog in our street chased

aaien dat kleine rode poesje die grote hond uit onze straat verjoeg

Gert in the sea drown is swimming

Gert in zee verzuipen zwemt

In order to demonstrate that one of the main problems students encounter when writing is to guess what sentences their strings blend into, and to test the hypothesis that Sentence Construction students will write more genuine sentences than School Grammar and Direct Writing students will, a sister was born to T-unit ratio. She is called sentence ratio. Sentence ratio concerns itself of necessity with flawless strings only, checks whether these are demarcated in one way or the other as 'sentence' and whether this sentence is a sentence indeed. The rules that stated what was to be considered a flawless string and what was not also guided raters in their judgement on sentences. Sentences, however, can be flawed in a way that T-units cannot: there may be no end to them. In sentence analysis, this error type was added to the instructions for raters. The sentences to be analyzed were whatever students had earmarked as such, however unconventionally. In this study, students numbered sentences or started each new sentence on a new line, put asterisks or dashes in front of new sentences, either began new sentences with enormous capitals or ended them with pitch-black full stops of three or four millimeter in diameter or, sophisticatedly though bleakly and uncreatively, just used capitals and full stops. Sentence ratio is the ratio of all words in studentwise demarcated faultless sentences and all words in the essay. Ideally it equals T-unit ratio, signifying that all strings are correctly blended into sentences (boundaries not necessarily coinciding). As the situation is far from ideal, however, T-unit ratio will generally tower above sentence ratio, thus neatly lending lustre to the problem writing genuine sentences poses.

Sentence ratio is accompanied by two other measures, both closely related to two of the measures accompanying T-unit ratio: the mean length of sentences

6. Or two. Instruction was tuned to sentence types.

and the ratio of noun phrases per sentence. These measures, functioning just like their T-unit pendants, are the finishing touches to the picture sentence ratio paints.

Scores on sentence awareness, then, are assessed by two discrete point tests and three measures. The tests are

- Recognizing Sentences;
- Constructing Sentences.

The measures are

- sentence ratio;
- mean sentence length;
- ratio of noun phrases per sentence.

As the latter two are merely meant to explore details, the first measure, sentence ratio, is the one measure that will be used to test the hypothesis on sentence awareness in writing.

Set at essay (5.1) the trio yield a sentence ratio score of .94, a mean sentence length of 9.05 and a ratio of noun phrases per sentence of .93. Set at (5.2) they yield a sentence ratio score of 0.00, a mean sentence length of 0.00 and a ratio of noun phrases per sentence of 0.00.

5.5 Conciseness

Sentence Construction students will be better able to tightly pack noun phrases, says one of the hypotheses. There is a test for this one, too. This one has students combine two or three kernel sentences into one sentence. The number of words students were allowed to use for this resulting sentence was limited. Reliability of this test, Informing Succinctly, was satisfactory (estimated alpha for 10 items: .86); parallel test versions were composed in the usual way. Here are two examples of items in the test.

Tell in one sentence what it says in two sentences. Put each word of that sentence on a separate line.
(Vertel in één zin wat hier in twee zinnen staat. Schrijf elk woord van die zin op een aparte streep.)

Rob kicked his friend.
His friend is small.

Rob schoot zijn vriendje.
Zijn vriendje is klein.

Tell in one sentence what it says in three sentences. Put each word of that sentence on a separate line.

I have eight turtles.
They are white.
My turtles are noisy.

Ik heb acht tortelduiven.
Ze zijn wit.
Mijn tortelduiven zijn rumoerig.

To test whether Sentence Construction students indeed employ complexly structured noun phrases in their writing more frequently than do School Grammar and Direct Writing students noun phrases in all student writing are categorized into simple and complex ones. The categorization is based on weighing the amount of information a noun phrases supplies and its structural complexity. Complex noun phrases are the ones containing either at least one adjective (*her oldest brother; two big yellow tulips; a fab sweet cookie* (Du: *een onwijs lekker koekje*), a prepositional phrase (*a doll in a shop window*) or an embedded phrase (*two boys who are able fighters*). Noun phrases with coordinated nouns both of which are specified by at least one determiner (not an article) get classified as complex as well, as were noun phrases combining any of these aspects. All other noun phrases were classified as simple noun phrases, including proper names. Book titles, quotes and enumerations came under this heading also. Pronouns were not classified at all. This categorization enables computation of the ratio of complex noun phrases and the total number of noun phrases in the essay.

Essay (5.1) scores .20 on this measure, (5.2) .14.

All in all, this amounts to one test and one measure for consiseness. They are:

- Informing Succinctly (discrete point test);
- complex noun phrase ratio.

5.6 Overall writing quality

The final hypothesis states that the Sentence Construction students will write impromptu essays of an overall quality superior to those of the Direct Writing students and these in turn will be superior in overall quality to those of the School Grammar students. Overall quality is the only non-syntactic aspect of writing involved in this study. Information on this aspect of student writing is gathered by holistically rating essays. What is rated exactly is unknown. Some say it is chiefly length of essays (Nold & Freedman 1977), some say it is length plus freedom from spelling errors (Stewart & Grobe 1979), some say it is cohesion and coherence, some say it is either, some say it is neither (for an exhaustive overview, see Spiegel & Fitzgerald 1990). Vocabulary has at least once been mentioned as an important factor to explain differences in quality (Grobe 1981). Headed nominal constructions might also account for these differences (Potter 1967). Syntactic complexity of whatever kind seems to be a likely candidate anyhow to account for differences in overall writing quality. Sentence combining students almost without exception outperformed non-sentence combining students on this variable (O'Hare 1972; Kerek et al. 1980; Strong 1986; the best-known exception is, of course, Mellon 1969; also cf. chapter 2.6). Perhaps, though, syntactic complexity covers something deeper, as is suggested by a meta-analysis of results of the overall quality ratings in sentence combining and sentence construction experiments (Hillocks 1986, also

cf. chapter 2.6). Then again, there might be a limit to the salutary effects of syntactic complexity on overall quality, and quality of some writing may be so miserable that not even sophisticated syntax will patch it up (Neilsen & Piché 1981). Obviously, researchers do not agree when it comes to penetrating good quality writing. Be this as it is, there still seems to be some sort of common sense agreement on what is good or bad quality writing. At least, raters do agree up to a certain limit.

The pilot study on the holistic rating of essays was conducted to try out whether a protocol was necessary. There was no serious check on inter-rater reliability. Nine raters and 450 randomly chosen essays, 150 for each rater, were involved. The essays all originated from the experiment. All raters were familiar with the kind of texts nine- to ten-year-olds compose, eight of them had in one way or another ever been involved in teaching in primary school. They first had to sift essays in three categories. Essays could be remarkably wretched, remarkably good or not remarkable at all. Next, raters had to sift the not-remarkable-at-all essays again, again in three categories. Essays then could be better or worse than not remarkable or still be not remarkable. As it turned out, this procedure was fine, so, in the following runs of holistic scoring it was used over and over.

5.7 Summary

In this chapter, construction, try out and revision of the measuring instruments for the dependent variables are reported upon. Four of fourteen tests, clustering the surviving 215 of 1123 items, met the criteria set for difficulty (*p*-value) and reliability (coefficient alpha). Data on these tests are presented in table 5.1. These four tests yield information on academic linguistic knowledge (Recognizing Verbs), on sentence awareness (Recognizing Sentences and Constructing Sentences) and on conciseness (Informing Succinctly).

To measure syntactic aspects of writing, eight instruments have been developed. These assemble information on writing fluency, sentence awareness and conciseness in student prose. Non-syntactic aspects of writing are assessed by having juries holistically rate essays without any instruction to speak of.

Table 5.2 shows what method (discrete point test, guided or non-guided writing) is used to measure what variable (writing fluency, academic linguistic knowledge, sentence awareness, conciseness or overall writing quality) and what are the instruments that do so. Notice the striking resemblance this set-up bears to a Swiss cheese. Moreover, note that the method 'guided writing' coincides completely with Guided Description. Table 5.3 summarizes all hypotheses to be put to the test.

Table 5.1. Discrete point tests; variables these gather information on and mean difficulty or p-value, mean standard deviation and mean reliability for five parallel test versions

Name of test	Variable	Number of items per version	Estimated mean p-val. per version	Estimated mean sd per version	Estimated mean reliability per version
Recognizing verbs	academic linguistic knowledge	14	.53	.24	.76
Recognizing sentences	sentence awareness	9	.50	.24	.78
Constructing sentences	sentence awareness	10	.90	.24	.78
Informing succinctly	conciseness	10	.43	.34	.87

Table 5.2. Instruments (tests and measures) used to measure five variables via three methods

Variable	Tests and measures		
	Discrete point test	Guided writing: Guided Description	Non-guided writing: three impromptu essays
writing fluency		main measure - T-unit ratio exploratory measures - noun phrases per T-unit - mean T-unit length	main measure - T-unit ratio exploratory measures - noun phrases per T-unit - mean T-unit length - run-on ratio
academic linguistic knowledge	- Recognizing Verbs		
sentence awareness	- Recognizing Sentences - Constructing Sentences	main measure - sentence ratio exploratory measures - noun phrases per sentence - mean sentence length	main measure - sentence ratio exploratory measures - noun phrases per sentence - mean sentence length
conciseness	- Informing Succinctly	- complex noun phrase ratio	- complex noun phrase ratio
overall writing skill			- quality score

Table 5.3. Hypotheses (Sentence Construction=SC, School Grammar=SG, Direct Writing=DW)

Variable	Test or measure	Hypothesis
writing fluency	T-unit ratio	SC = SG = DW $\geq$.80
academic linguistic knowledge	Recognizing Verbs	SC = SG > DW
sentence awareness	Recognizing Sentences Constructing Sentences sentence ratio	SC > SG = DW
conciseness	Informing Succinctly ratio of complex noun phrases	SC > SG = DW
overall writing quality	overall quality score	SC > DW > SG

ON THE ROAD TO NEVER-NEVER LAND: EXPERIMENTING

'... But much of our education we got without knowing the why or the wherefore at all. That is perhaps the best explanation of the sixth sense: just knowing a thing without knowing why or how you know it ... Altogether, as a grown-up bird I've often astonished myself at what a lot I know - and how I know it I couldn't tell you to save my life. But I'm convinced that a great deal of the most important part of my education came to me, as it does to all of us, during that time when we lie in the nest with our eyes closed, trying to guess, by hearing and smelling and feeling, what the world is going to look like.'

*Pippinella the canary to Doctor Doolittle in Hugh Lofting,
Doctor Doolittle's caravan.*

This chapter deals with the execution of the experiment and all sorts of matters connected with it. Method is described in section 1. Relevant information on matters such as subjects, phasing of the experiment and scheduling of the test periods is to be found there. Section 2 informally sketches the daily labours of the three experimenters and their students and formally reports upon matters such as protocol adherence, time on task, students' progress and interference of the regular curriculum¹. In section 3 attention will be paid to the psychometric quality of the collected data. It presents reliability of tests and measures for the five dependent variables. Section 4 concerns preliminaries to and plans for analyses of data. Section 5 summarizes this chapter.

6.1 Method

Subjects

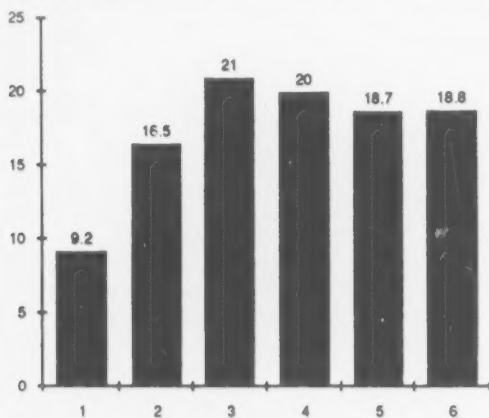
The number of subjects needed per treatment was computed to be 52 (moderate effect in univariate analysis of variance; power: .80; $\alpha=.05^2$), so six schools with an average of 30 students per class had to be recruited. A selection of the

-
1. A detailed report on the data this quartet yielded is to be found in Van de Gein (1990).
 2. Cohen, J. (1977) *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

schools involved in the pilot study were asked to participate. Five of these, with a total of some 130 students, were willing to do so. The sixth to participate was a school that had been working with try out materials of *Grammatica in Balans*³ for several years. There were 38 students in its sixth form and so, by their participation, the number of subjects was complete.

Three of the schools participating in the study were situated in cities in the north-western part of Holland (Blokker, Castricum and Hoom), three were situated in cities in the mid-western part of Holland (Dordrecht, The Hague and Vlaardingen). As the study into the effects of grammar instruction upon aspects of writing of primary school students was a true experiment, there was no need to bother overmuch over differences between schools. However, a glance at the scores per school on the standardized reading test *Schriftelijke Opdrachten 4* (Eng.: *Written Directions 4*) (Brus et al. 1979) (see figure 6.1) reveals there were differences (two-way fixed effect analysis of variance; $df=5$; $F=9.90$; $p=.00$).

Figure 6.1. Reading skill of students per school (1-6); number of correct answers



Students were assigned randomly to treatments, at every school resulting in three treatment groups. Experimenters were systematically allocated to schools and treatments in such a way that each experimenter taught the one treatment on two schools, the other on another two schools and the third on the next two schools. In table 6.1 assignment of experimenters to treatments per school is presented. Number of students per school, experimenter and treatment is presented, too.

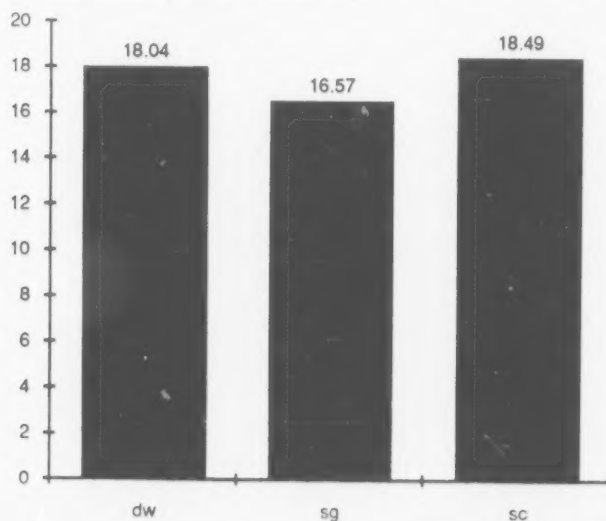
3. *Grammatica in Balans* is the grammar course the Sentence Construction programme was distilled from (also consult chapter 4.4).

Table 6.1. Assignment of experimenters (A, B, C) to treatments; number of students (n)

	TREATMENT		
	Direct Writing	School Grammar	Sentence Construction
school 1 n = 21	A n = 6	B n = 7	C n = 8
school 2 n = 30	A n = 10	B n = 10	C n = 10
school 3 n = 18	C n = 6	A n = 6	B n = 6
school 4 n = 38	C n = 12	A n = 13	B n = 13
school 5 n = 16	B n = 6	C n = 5	A n = 5
school 6 n = 42	B n = 14	C n = 14	A n = 14
Total	n = 54	n = 55	n = 56

Not surprisingly, differences in reading skill of students per treatment, presented in figure 6.2, are negligible (two-way analysis of variance; $df=2$; $F=1.53$; $p=.22$).

Figure 6.2. Reading skill of students per condition (Direct Writing=DW, School Grammar=SG, Sentence Construction=SC); number of correct answers



Phasing of the experiment; phasing of the test periods

The experiment was planned in two phases, actually resulting in two separate experiments. One is the first experiment or the experiment proper, carried out as described above (see table 6.1). The other is the extended experiment. The extended experiment comprises the experiment proper plus a follow-up treatment. Follow-up treatment was given in order to, as it were, 'repair' the forms that had participated in the experiment to some sort of common level; eventually, one of the three experimental courses had to be taught to all students in one class. Clearly, then, repairing classes was a matter of service in the first place: class teachers, not research questions, determined what the follow-up course had to be: Direct Writing, School Grammar or Sentence Construction. As a matter of course there always was one group at each school that had reached this 'common level' already. The questions pertaining to the effects of the follow-up treatment (cf. chapter 3.2) determined what had to be the follow-up treatment for this group: if the first course had been Direct Writing, its follow-up was Sentence Construction, if the first had been Sentence Construction or School Grammar, then it was Direct Writing.

Though, in a way, this 'repair treatment', as it came to be known, simply meant running the first experiment over, its experimental design got slightly tainted, as can be seen in table 6.2. The most important taint is the contamination of school and treatment (school 5 x School Grammar treatment). Also, not all students were enrolled in an ordering of treatments that would help answer the questions at stake (again consult table 6.2; students not at all enrolled in the Direct Writing treatment are a case in point). Finally, it seemed safest to study follow-up effects of students in same schools only. All this means the number of students involved in the follow-up studies is far less than 165. As table 6.3 shows, 99 students are involved in the first study, and only as few as 60 in the second. As is clear from table 6.3, both studies involve students enrolled in the Sentence Construction course in the first experiment and in Direct Writing in the second. Note that, due to the 'same schools'-requirement, not all students taught to construct sentences first and to construct texts next are involved in both studies, though. In the first study, effects of this ordering of treatments will be compared to effects of courses ordered reversely. The second study compares effects of the order Sentence Construction-Direct Writing to effects of the order School Grammar-Direct Writing. It should be clear by now that neither of the follow-up studies are to be likened to the hard-boiled effect study the experiment proper was. Rather, both extended experiments, i.e.: experiment proper plus follow-up study 1 or 2, are delicate explorations indeed, to be engaged in very carefully.

The experiment proper was run during the first half of the school year. It started right after summer vacations in week 35 (August) and ran on until week 51 (December). There was a break of one week. The second experiment took up the other half of the school year. It started right after Christmas holidays in week 2. This experiment was wound up in week 17 (May). There were two one week breaks in this experiment.

In terms of research, there was birth rather than mortality of students as at the end of the study there were six students more than there were at the start. During the experiment proper no students were lost. The second experiment started out with two students less but ended up with eight more. Test scores of the latter students have not been analyzed.

Table 6.2. Second experiment. Assignment of experimenters (A, B, C) to follow-up treatments; number of students per cell (n)

1st treatment-2nd treatment	School						Total
	1	2	3	4	5	6	
Direct Writing-School Grammar	-	-	-	-	B n=6	-	6
Direct Writing-Sentence Construction	A n=6	C* n=10	C n=6	C n=12	-	B n=16	50
School Grammar-Direct Writing	-	B n=10	A n=6	-	C n=5	C n=14	35
School Grammar-Sentence Constr.	B n=7	-	-	A n=13	-	-	20
Sentence Construction-Direct Writing	C n=9	A* n=9	B n=6	B n=13	-	A n=17	54
Sentence Constr.-School Grammar	-	-	-	-	A n=5	-	5

* As experimenter A threatened to succumb under gigantic loads of correction work in the second experiment, she and experimenter C exchanged groups at school 2.

Table 6.3. Extended experiment. Number of subjects involved per study

school	First study		Second study	
	Sentence Constr. followed by Direct Writing	Direct Writing followed by Sentence Constr.	Sentence Constr. followed by Direct Writing	School Grammar followed by Direct Writing
1	6	8	--	--
2	10	10	10	10
3	6	6	6	6
4	12	13	--	--
5	--	--	--	--
6	14	14	14	14
total	48	51	30	30

There were three test periods in the experiment proper: a pre, a mid and a post test. In the second experiment there were two more test periods: a mid and a

post test. Test periods came in intervals of six or seven weeks. Each test period consisted of three sittings, to be spread over three (consecutive) days. All sittings were set up the same, as can be seen from table 6.4.

Table 6.4. Set up of all five test periods in sittings

Sit	Test/assignment	Duration in minutes (instr. included)	Variable
1	Recognizing Verbs	15	academic linguistic knowledge
	Constructing Sentences	25	sentence awareness
	Description*	20	writing fluency
			sentence awareness
2	Informing Succinctly	30	conciseness
	Narration*	30	writing fluency
			sentence awareness
			conciseness
3	Recognizing Sentences	15	sentence awareness
	Guided Description	25	writing fluency
			sentence awareness
			conciseness
	Exposition*	20	writing fluency
			sentence awareness
			conciseness
			overall writing quality

* To control for topic effects, topics were allocated to students and test periods systematically (also cf. chapter 5.1).

6.2 Work in progress: quality matters

General information

A total of 936 lessons were given, 312 by each experimenter. Experimenters travelled over 500 hours each by train, bus and taxi to get to the schools, often hurriedly. Due to a railway power failure and delays some lessons started later than scheduled. Students had to do without Miss as one of the Misses once took a train in the wrong direction. Once a taxi crashed, but this did not keep experimenters from working, shaken though they were. As a matter of fact, almost nothing kept them from working, be it flu, back-aches, glass-cut foreheads or broken bones. Halfway through the second experiment, though, one experimenter fell ill far too seriously to be able to keep on working. By joining groups the other two experimenters were able to take care of her groups on schools 1 to 5; a recently graduated educational linguist with teaching experience was found willing to substitute her on the sixth school.

No experimenter was confronted with serious problems concerning the management of students. Each, though, had one group they particularly liked and one they particularly disliked, all, for that matter, at different schools. The way

groups were assembled (randomly!) turned out to be the cause for the dislike experimenters took to groups. As one class teacher put it: 'I am not surprised. You should have asked me to form groups. These kids simply shouldn't be together in one such a small group! You're asking for trouble.'

In the experiment all things that can possibly happen when, in a period of hardly 30 weeks over 900 classes are taught, actually happened. Students entered the classroom fiercely fighting, in neighbouring classrooms thirty noisy students were taught a very noisy song, loud music was to be heard all over the place in preparation of Mardi Gras, Miss forgot to take along students' workbooks and once a silent row of gnomes passed through the classroom. Still, with relatively little effort experimenters were able to keep tempo of groups up. Or down. Except for the penultimate lesson in the Direct Writing course, all material was covered.

Protocol adherence

As it turned out to be impossible to carry out the experiment double blindly, teacher bias threatened its results, albeit experimenters were allocated systematically to treatments (also cf. tables 6.1 and 6.2). So, behaviour of experimenters was tightly restricted by their using protocols. Though training in using these protocols was harsh and successful, data were gathered on Miss' doings in class all the same: all lessons were audiotaped (also cf. chapter 3.4). For each experiment, per treatment per experimenter audiotapes of four random lessons out of every six or seven were sampled in order to gather information on behaviour of experimenters. There were 62 tapes included in the sample, yielding 112 fragments of class teaching, totalling about 40 hours⁴. Tapes were compared to the protocols by ticking correspondences, digressions, omissions and additions. Analysis concentrated on phasing of lessons, on terminology used and on the scrupulousness with which the experimenters executed some almost ritual activities at the heart of treatments.

Protocol adherence was found to be satisfactory. Adherence to phasing and terminology was even more than satisfactory. Concerning the ritual activities, one of the experimenters turned out to be rather less strict than the others, asking questions that focused on irrelevant matters and adding faulty detail. Still, even this slightly self-willed experimenter held on tightly to the protocol. There was no evidence whatsoever that experimenters put more effort in one or the other of the treatments. Though the life of an experimenter is certainly not to be compared to the life of an average teacher, the experimenters were very true teachers in one aspect: they wanted to teach the students certain things and they did their utmost to do so, manoeuvring skilfully within tightly set bounds.

4. Of course, there was much more worth while on these tapes, such as an angry experimenter hauling a group over the coals, students of other forms asking whether somebody had spotted a pink bathing suit, curious class teachers who came to have a look, students serving coffee or tea to an evidently delighted experimenter, experimenters reading stories to their groups and of course the odd student, unobserved by Miss, whispering 'hello' into the microphone of the tape recorder.

Time on task

In order to check whether there were systematic differences between conditions in the amount of time spent on course activities, data on students' time on task were gathered (also cf. chapter 3.5). This was done by observing two lessons per experiment per treatment at all schools, amounting to 72 lessons. Four randomly chosen students were observed in turn for one minute, the observer scoring the student's most characteristic activity every twentieth second. In this way, in the first experiment 4776 activities were observed, in the second 4642. Results for both experiments are presented in table 6.5.

On the average, classes took just under one hour. In the first experiment, there were no differences between treatments in time spent on treatment activities, in the second experiment there were. This difference, however, is due to contamination of school x treatment. In both experiments, Sentence Construction students received least instruction and spent most time on unsupervised seat work. This tallies with the character of the course (cf. chapter 4.4, esp. the examples taken from the course). However, Direct Writing students in the second experiment worked longest of all: longer than their Direct Writing colleagues in the first experiment and longer than both School Grammar or Sentence Construction students in either experiment. On the whole some ten minutes per class were spent on other than true course activities.

Table 6.5. Mean time on task of students per class per treatment in minutes. (Direct Writing=DW, School Grammar=SG, Sentence Construction=SC)

Treatment:	experiment proper			second experiment		
	DW	SG	SC	DW	SG	SC
No. of activities:	1601	1533	1642	2497	424	1721
Instruction	23	24	19	21	20	17
- to whole class	21	22	15	19	19	15
- individual	2	2	4	2	1	2
Seatwork	23	23	27	29	19	27
- supervised	3	0	0	2	0	1
- unsupervised	20	23	27	27	19	26
Spent on treatment	46	47	46	50	39	44
Other activities	9	10	11	7	6	13
Total	55	57	57	57	45	57

As table 6.5 shows, in the first experiment for each treatment over 80% of time was spent on treatment activities. So, it seems safe to conclude that, overall, students in the first experiment spent enough time on subject material. In the second experiment this only goes for Direct Writing and School Grammar students, Sentence Construction students spending 77% of class time on treatment activities. In other words, all students except those participating in the Sentence

Construction condition in the second experiment spent enough time on course material⁵.

To compute systematic differences between treatments concerning time spent on treatment activities, an analysis of variance was run. In order to pin down as precisely as possible what source (treatment, experimenter or school) caused systematic differences, a three-way analysis of variance was called for: all three sources come as main effects in such an analysis. Results are presented in table 6.6.⁶

Table 6.6. Experiment proper. Results of three-way analyses of variance for the effects of treatment, experimenter and school on time spent on treatment activities

Source	dF	F-ratio	p
Treatment	2	.11	.90
Experimenter	2	.72	.50
School	5	1.56	.21

This table shows that 'school', not treatment or experimenter, is the source that causes systematic differences (in order to increase power, alpha was set at .25). Detailed analysis of data revealed that the schools responsible for this effect are numbers 4, 5 and 6, i.e. the largest two (4 and 6) and the smallest. Apparently, (and probably not surprisingly) number of students per group determined the amount of time spent on treatment activities. The larger the group, the more ado; in the smallest of groups, the tempo was highest.

Students' progress

Experimental effects might well be influenced by systematic differences between students' progress per treatment group. After all, poor performance of students might have been caused by their not attending lessons, literally or figuratively, or by their not having done enough work in their tradebooks, or from not having done well enough. Thus experimenters kept track of students' progress (also cf. chapter 3.6).

5. It turned out hard to decide how come. Rather meticulous analysis of data suggested that the 'large' number of students in the (joined!) Sentence Construction groups (in the second experiment thirteen for school 1 and twentyfive for school 4, cf. table 6.2) caused the two surviving experimenters some minor managing problems. Also, students had grown older: there is the remotest of suggestions that the course might have been a tiny bit too simple, now and then, causing students to finish their work within the hour, leaving them free to do something else. If this surmise is correct then, of course, students spent enough time on material but there simply was not enough of it.

6. As the far from perfect experimental design obscures results, no statistical analyses on the data of the second experiment are presented here. However, data freaks might want to consult Van de Gein 1990.

Progress of Direct Writing students was checked by keeping track of their presence and of the amount of writing done. No fault was found, there. Systematical differences between the School Grammar and the Sentence Construction courses only will be discussed.

Subject material of both the School Grammar and the Sentence Construction course came in three sorts: material to be covered jointly in class (C-exercises), material to be covered by students individually (I-exercises) and extra material (E-exercises). Both C- and I-exercises had to be covered and mastered (mastery norm: 80%). Of the E-exercises that students had worked on, 50% had to be correct.

There were 127,963 data (ticks or hyphens) available. Analysis was performed on a sample of 1260 of these, randomly chosen from each of the three sorts of exercises. Results are presented in table 6.7 for both experiments. As this table shows, both School Grammar and Sentence Construction students in both experiments easily met the set norms⁷. So, though statistical analysis (chi-square) revealed there was a systematic difference between treatments for E-exercises only (alpha set at .25; first experiment, $p = .12$; second experiment: cf. note 6), there seems to be no need to worry about students' progress.

Table 6.7. Percentage of covered and mastered of all subject matter per treatment

	to be covered jointly		to be covered alone		extra exercises	
	covered	mastered	covered	mastered	covered	mastered
First experiment						
School Grammar	98	97	99	94	70	63
Sentence Constr.	98	95	98	97	74	63
Second experiment						
School Grammar	98	98	100	100	90	86
Sentence Constr.	92	92	95	92	82	72

Regular curriculum

The regular class teachers had been handed out shadow books to work with during the time of the experiment. In these shadow books (extra copies of the

7. One of the experimenters carried out a qualitative analysis of the way students in Sentence Construction groups in two of the participating schools came to terms with the material presented to them. Her results corroborate findings reported here (Braams, 1989).

language arts material used at the school) all subject material to be avoided was crossed out. Shadow books were meant to take down notes in, too. In order to regularly check the teacher's accuracy for each teacher there were two extra copies of the shadow book. These were exchanged every now and again (also cf. chapter 3.7).

Class teachers very carefully kept track of their language arts activities. Some of the shadow books were amply annotated. As it turned out, students of school 2 once punctuated a text and once constructed two word sentences. During the second experiment, school 5 students occasionally were confronted with School Grammar material in its genuine *Jouw Taal Mijn Taal*-format. At school 6 several of the students once wrote a Punch and Judy play and one or two contributed to the school newspaper. Students involved in these activities, though, can hardly be said to have been involved in extra instruction. This warrants the conclusion that no systematic interference of the regular language arts program has taken place.

Experimenter effects

In the research design employed in this particular study, two-way analyses of variance yield experimenter effects as two-way interaction effects. So, effects of treatment $\times$ school were analyzed by running two-way analyses of variance over all tests and measures for all test periods, totalling 72 analyses. As exactly one and one only of these was found to be significant statistically (alpha set at .05) it seems safe to conclude that the effects of the experiment will not be soiled by experimenter effects.

6.3 Quality of data (with Mindert Eiting)

Trained test assistants equipped with detailed protocols directed all test periods. They took notes on irregularities during test taking and registered remarkable student questions⁸. Though extreme care was taken to collect a complete set of data of all students, it turned out to be impossible to have no missing data at all. In order for scorings to be performed blindly, all incoming material was coded. Essays were typed and prepared for analysis (for an example: see translations of essays in chapter 5.2).

Quality of the four discrete point tests is presented in table 6.8. Though for some less than hoped for --especially Recognizing Verbs and Recognizing Sentences at the time of the pre test (1st tp), as well as Recognizing Sentences at the fifth test period-- overall, reliability of tests was satisfactory. Declining reliability of the tests for Recognizing Sentences and Constructing Sentences is

8. Such as 'Can I have an extra sheet to write a draft, first?' asked by about one thirds in each class when starting on the impromptu essay.

probably due to ceiling effects. Be that as it may, no tests have to be dropped in an analysis of experimental effects.

Table 6.8. Reliability (coefficient alpha) of discrete point tests for all five test periods (tp); m = number of items

test	1st tp	2nd tp	3rd tp	4th tp	5th tp
Recognizing Verbs m = 14	.39	.65	.62	.72	.86
Recognizing Sentences m = 9	.57	.61	.69	.70	.50
Constructing Sentences m = 10	.66	.73	.71	.65	.60
Informing Succinctly m = 10	.78	.88	.89	.89	.90

Nine raters, all students of or graduates in linguistics, performed the syntactic analyses of the guided and non-guided open writing tests. All essay ratings were performed by pairs of raters, according to the Van den Bergh and Eiting (1989)⁹ design. So all 2475, coded with random numbers, were shuffled and divided randomly over a total of nine sets, each to be rated twice. Mean individual reliability of raters, estimated according to Van den Bergh and Eiting (*ibid.*) turned out to be .72, reliability ranging from .64 to .84.

Classical (non-rater) reliability of the sum of four measures (three essay genres and the guided writing test per period) was estimated in a model for congeneric items¹⁰ utilizing the Fabin 2 method¹¹. Results are presented in table 6.9. Apart from sentence ratio, mean sentence length and noun phrases per sentence, syntactic measures were unreliable (mean reliability not even amounting to .50). It seems likely this is due to lack of systematic variance.

Overall quality of essays was rated in two runs. The first run was meant to establish what was the correlation of overall quality and syntactic quality (defined by sentence ratio) as the co-author of this section deemed it wellnigh possible that this correlation, corrected for attenuation, would amount to .95 or thereabouts. If this were the case, separate ratings of quality would be rather superfluous. A completely randomized set of 25 percent of all essays written in the experiment, an equal amount of randomly chosen descriptions, narratives

9. Bergh, H. van den & Eiting, M.H. (1989). A method of estimating rater reliability. *Journal of Educational Measurement*, 26, 29-40.

10. Fleishman, I. & Benson, J. (1987). Using LISREL to evaluate measurement models and scale reliability. *Educational Psychological Measurement*, 47, 925-939.

11. Hagglund, G. (1982). Factor analysis by instrumental variables methods. *Psychometrika*, 47, 209-222.

and expositions, was rated in juries of four. All were experienced primary school teachers. Jury reliability was estimated to be .82; satisfactory, that is.

Table 6.9. Reliability (Fabin) of measures for student writing per test period (tp)

	1st tp	2nd tp	3rd tp	4th tp	5th tp
writing fluency					
mean measure	.29	.48	*	.42	.46
T-unit ratio					
exploratory measures					
mean T-unit length	.38	.38	.48	.52	.34
run-on ratio#	.45	.41	.37	.51	.34
ratio of noun phrases per T-unit	.26	.48	*	.39	.46
sentence awareness					
mean measure	.63	.72	.72	.69	.71
sentence ratio					
exploratory measures					
mean sentence length	.42	.48	.49	.66	.74
ratio of noun phrase per sentence	.62	.67	.71	.65	.69
conclusiveness					
complex noun phrase ratio	*	*	*	*	*
overall writing quality					
overall quality score#	.60	.70	.69	.66	.70

* Reliability cannot be computed due to negative parameter estimates.

Reliability computed over three non-guided writing tests only.

Subsequently, the correlation between overall quality, corrected for rater attenuation, and syntactic quality, corrected similarly, was estimated (utilizing a Lisrel model for congeneric measurements). The estimate proved to be .25¹². Hence, the remaining 75 percent of essays were to be rated likewise.

Overall quality ratings of the remaining 75 percent of essays were performed by twelve experienced primary school teachers in three juries of four. Reliability of these juries ranged from .80 to .83; mean reliability of all five juries thus amounts to .81. Classical (non-rater) reliability of quality ratings, the sum of three essay genres per period, is also presented in table 6.9.

As no analysis should be run over psychometrically unreliable measures or will be run over the exploratory ones, experimental effects in actual junior writing will be computed by sentence ratio and overall quality scores, these being the only ones more or less well above .60.

12. As we had turned our difference of opinion into a real bet, with a wafer hat for stake, I had taken special care to make this an extra bitter, large and dry pill to swallow.

6.4 Prelude to data analysis

In order to gain insight in the overall relationship between scores, correlations were computed for data collected at all five test periods. However, results will only be presented here for data collected at the third test period, i.e. the last of the experiment proper. Reliability of all tests and most measures just passes at this period (cf. tables 6.8 and 6.9). A prior word of warning is due. Low or incomputable reliability of scores (such as, e.g., the scores on T-unit ratio and run-on ratio) obscures correlations. Dependency of scores causes correlations to be over-estimated artificially. In this study, such dependency abounds: between, e.g., T-unit and sentence ratio, fractions of both having identical denominators, viz. total number of words in the essay; between, e.g., sentence ratio and mean sentence length, fractions of both having identical numerators, viz. total number of words in student-wise demarcated fully blown sentences; or, for that matter, between all measures, distilling nine 'separate' scores out of one impromptu essay (for an illustration, consult table 4.2 and the last column of table 6.4). Hence, as interpreting these data is hazardous, blatantly weird relationships will be dealt with only.

In the first run, correlations were computed between all four tests and all nine measures, summed over all four writing tests. A complete table presenting all correlations is to be found in appendix 2. It seems there are two pairs of measures far too intimately related: correlation between T-unit ratio and the ratio of noun phrases per T-units amounts to .94; correlation of sentence ratio and ratio of noun phrases per sentence amounts to .99. This comes as a mild surprise, for it means as much as their being identical twins. How come? The twin sisters meant to measure what was the amount of noun phrases coming in full clauses (or T-units) and in flawless sentences in relation to the total amount of noun phrases. Both obviously assume silently that young writers are either apt to thread noun phrases to make a list or to exploit noun phrases to carry content. So, in first instance both measures seem inclined to distinguish between the diverse types of erroneous strings and sentences heaped together indiscriminately by T-unit ratio and sentence ratio (also cf. chapter 5.2 and 5.4). How come neither fulfils its promise?

Treading down the elegant tradition of patching up clauses is payed for, here, it seems. The one measure allows computing the ratio noun phrases in a T-unit, the other allows computing the ratio noun phrases in a sentence. However, in order for a noun phrase to be registered as part of a T-unit or sentence, the T-unit and sentence have to be absolutely flawless. Thus, both measures are intricately interwoven with the notion of T-unit and sentence operative in this study. By thus pervading these measures, T-unit ratio and sentence ratio turned both measures into their twins quite unthinkingly. In fact, flawed and flawless T-units as well as flawed and flawless sentences simply are measured twice over. Evidently, when one wants to explore the way writers frame content, one had better do so directly, e.g. by making an inventory of the structures noun

phrases come in, irrespective of the shortcomings in the way they appear in writing. As results of the one (T-unit ratio and sentence ratio respectively) imply results of the other (ratio of noun phrases per T-unit and sentence respectively), half of both twins might as well be dropped in presentation of data. As T-unit and sentence ratio are the more important issues of this study, and main measures at that, the noun phrase twin sisters of both will be dropped.

Connoisseurs are kindly requested to note that the correlation of both sentence awareness tests, Recognizing and Constructing Sentences, does not indicate these measure exactly the same thing ($r = .49$). These tests, then, will not be summed.

In the second run, correlations between measures were taken into consideration in more detail. This was done by computing correlations of measures separately for all writing tests, leaving out, of course, the two drop outs. A table presenting these correlations at the third test period only is to be found in appendix 3. As a matter of course, there is another twin to be found, there: the correlation between T-unit ratio and sentence ratio for Guided Description. This one amounts to a firm .99. In sentence ratio flawlessness as well as demarcation of sentences is evaluated. When writing under guidance, students only write one sentence at the time (cf. chapter 5.2) so performance on this test is almost exclusively determined by flawlessness of sentences. More often than not in Guided Description these sentences coincide completely with the syntactic strings that are involved in the computation of T-unit ratio. Small wonder, then, that T-unit ratio and sentence ratio are duplicates. The one to be dropped in presentation is T-unit ratio, as sentence ratio will disclose whether students indeed wrote sentences when they were told to do so. Dropping T-unit ratio for presentation of data on this test means dropping mean T-unit length as well.

Analysis of the experimental effects will be performed by computing standardized contrasts between groups: differences between groups on post test scores will be expressed by relating these to performance at the time of the pre test. Such differences between group means (or: contrasts) on post tests are in general practice expressed in terms of the standard deviation of the pre test scores in all groups together. However, the standard deviation of the pre test is affected by sampling. Group means on the post test are affected likewise. Moreover, score distributions may deviate from normality. Also, pre test and post test scores are dependent. Hence, the statistical properties of the standardized contrasts may be unknown completely.

Standardized contrasts in this study, then, have been simulated by taking random samples (with replacement) from pre test and post test scores (according to group definition and size). In each trial standardized contrasts are

computed. This is repeated a thousand times¹³. Subsequently, the boundary contrasts enclosing 950 or 990 contrasts found are established. These define the 95 or 99 percent confidence intervals for the contrasts. Such intervals represent, as it were, the 'band width' within which the contrast varies. When zero lies within this interval, this indicates the contrast might as well have been zero. Such contrasts are best neglected.

Standardized contrasts plus interval boundaries will be presented for all four tests and for sentence ratio and overall writing quality. Contrasts on sentence ratios will be computed over both Guided Writing and over the summed ratios for Description, Narrative and Exposition. These will be summed as only effects that manifest themselves in all three will be seriously considered to be experimental effects (also cf. chapter 5.1; also note that correlations (cf. appendix 3) do not give rise to the suspicion that these scores and apples and oranges). Contrasts for overall writing quality will be computed over the summed essay scores as well.

6.5 Summary and conclusion

This chapter dealt with all sorts of matters concerning the experiment in progress. Actually, two experiments were carried out. One was the experiment proper, a true experiment involving one hundred and sixty five subjects, three experimenters and three treatments, viz. Direct Writing, School Grammar and Sentence Construction. The other was the second experiment, involving one hundred and seventy one subjects, three, eventually four experimenters, and the same three treatments. However, treatment and school were contaminated in this second experiment. The second experiment facilitated two follow-up studies. One contrasts effects of teaching students to construct sentences first and to compose texts next to the effects of courses taught vice versa. The other contrasts effects of School Grammar, followed by Direct Writing, to effects of Sentence Construction followed by Direct Writing.

In order to control nuisance variables data were collected on the quality of the work in progress. Protocol adherence of experimenters was satisfactory. Time on task observations brought out that all students except the Sentence Construction students in the second experiment spent enough time on course material. The latter spent a tiny bit too much time on matters not bearing on the course. Analysis of students' progress for School Grammar and Sentence Construction treatments showed students in both treatments to meet the set mastery norm easily. There was no systematic interference of the regular language arts programs. Two-way analyses of variance showed there were no experimenter

13. There was, of course, a computer programme doing this menial work. It is called after Baron Münchhausen. Mindert Eiting created it. He also assisted in this account of its doings.

effects. So, whatever the effects of the experiment proper will turn out to be, they certainly are experimental effects and nothing else.

There were five test periods, a pre, mid and post test for the first experiment and a mid and a post test for the second. Quality of discrete point tests was satisfactory. Quality of five out of nine measures, all five of them syntactic measures, was below standards.

Results of the computation of correlations brought to light that three measures have doubles. These are T-unit ratio (double: ratio of noun phrases per T-unit), sentence ratio (double: ratio of noun phrases per sentence) and, in Guided Description only, T-unit ratio again (double: sentence ratio).

Table 6.10 presents an overview of all five variables and their (i.e. of writing fluency and sentence awareness: main) tests and measures. Its final column says whether these will come in in the analyses of experimental effects. The latter will be analyzed by computing standardized contrasts over scores gathered by psychometrically reliable tests and measures only. The intervals that go with these indicate the degree of confidence to be put in results: forewarned forearmed.

Table 6.10. Variables, (main) tests and/or measures, to be used (or not) in analyses of experimental effects

Variable	Test/measure	To be used in analysis
writing fluency	T-unit ratio	no
academic linguistic knowledge	Recognizing Verbs	yes
sentence awareness	Recognizing Sentences	yes
	Constructing Sentences	yes
	sentence ratio	yes
conciseness	Informing Succinctly	yes
	complex noun phrase ratio	no
overall writing quality	overall quality score	yes

THE EXPERIMENTS: RESULTS

The study of English grammatical terminology had a negligible or even relatively harmful effect upon the correctness of children's writing.

Harris cited in Braddock (1963) and Hillocks (1986)

The study into the effects of grammar instruction upon writing of primary school students comprised two experiments: the experiment proper and the extended experiment. Results of both will be presented here. The experiment proper is first in focus. In this experiment 165 students were involved, randomly assigned to three experimental groups within schools. Per school, subjects were assigned randomly to one of three treatments, viz. to either a course in Direct Writing, a course in School Grammar or one on Sentence Construction. The extended experiment comprised two minor and exploratory studies into effects of specific ordering of treatments (also cf. chapter 3.2 and 6.1).

Student performance was measured five times over. There were pre, mid and post tests in the experiment proper. In the follow-up to this experiment, there were mid and post tests again. Results will be presented by way of means and standard deviations per condition per test period. Thus, reports on the experiment proper involve three test periods, viz. the pre test, mid test I and post test I. Reports on the extended experiment involve four test periods: the pre and post tests of the experiment proper and mid and post test II. As not all records of students are complete, the number of students reported upon will be presented in all tables.

Results will be presented per variable. So, section 1 deals with scores on writing fluency, section 2 with scores on academic linguistic knowledge and section 3 with scores on tests and measures for sentence awareness. The fourth section concerns conciseness and the fifth overall writing quality. All sections are patterned alike. First, results of the experiment proper will be presented. For those scores worthy of statistical analysis, effects of treatments will be presented in standardized contrasts next. Finally, results of the extended experiment will be presented. Section 6 summarizes the chapter.

Effects will be presented by contrasting pre and post test performance of the three experimental groups. These (standardized) contrasts will be presented in figures showing, as it were, poles and flags. The Sentence Construction treatment always functions as pole, School Grammar and Direct Writing treatments as flags. Flags waving right indicate positive effects (or: better results) of the latter two treatments as contrasted to Sentence Construction, flags waving left negative effects (or worse results). Black flags are School Grammar flags, spotted flags are put up for Direct Writing. Length of the flags indicates the size of the contrast. Some of the flags go with one or two asterisks. These relate to the estimated confidence intervals. One asterisk means the 95% confidence interval does not include zero, two asterisks mean the same for the 99% confidence interval.

As a final preliminary to reporting results and effects, means and standard deviations of length of texts (for the experiment proper only) are presented in table 7.1. There was no need to care for effects of topics within genre, as all topics were systematically allocated to students and test periods (cf. chapter 5.1).

Table 7.1. Experiment proper. Number of words per writing test; (rounded) means and standard deviation (sd)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
Guided Description	n = 162		n = 164		n = 165	
- Direct Writing	71	15	78	19	69	23
- School Grammar	72	14	78	17	63	17
- Sentence Construction	77	20	75	20	68	18
Description	n = 162		n = 164		n = 165	
- Direct Writing	32	16	37	21	33	18
- School Grammar	35	17	37	18	35	19
- Sentence Construction	34	18	37	20	35	19
Narrative	n = 159		n = 162		n = 162	
- Direct Writing	109	38	111	55	100	50
- School Grammar	103	39	109	48	109	65
- Sentence Construction	104	50	106	52	108	58
Exposition	n = 162		n = 164		n = 165	
- Direct Writing	54	28	51	35	50	30
- School Grammar	49	28	48	30	48	35
- Sentence Construction	48	28	56	36	50	32

As table 7.1 shows, each genre seems to have its own stable means and variance, not to be influenced by test period or treatment. Per test period, the

average student supplied a total of 250 words¹. The descriptive genre in its open format yielded the shortest texts, only half the length of its guided sister. It should be kept in mind, though, that the mean length of the guided description is a more or less artificial result: students had to write fourteen separate sentences (cf. chapter 5.1). Narrative on the average elicited the longest texts. Exposition lands somewhere in the middle between the other two genres. Large standard deviations (pointing out that there is no such student as an average student) are caused by tiny groups of devoted writers, one of these writing a narrative of some 400 words. Length of texts is, on the whole, modest, but then, so was the amount of time students were allowed to work on these (impromptu) essays.

7.1 Writing fluency

Tables 7.2 to 7.4 report upon writing fluency. Table 7.2 gives means and standard deviations for the main measure: T-unit ratio. Tables 7.3 and 7.4 do the same for both exploratory measures: mean T-unit length and run-on ratio.

The hypothesis on this aspect of writing has it that T-unit ratio would amount to at least .80 anytime anywhere: no treatment would, could or needed affect writing fluency. Low reliability of T-unit ratio (in all likelihood due to lack of variance) does not allow statistical analysis. Hence it is impossible to check whether the hypothesis holds. But then again, T-unit's lack of variance points out that, time and again, students score all alike. Doesn't this point out that, indeed, treatment does not affect it at all? It does; it doesn't. For this very reason, no results will be presented for the extended experiment.

As was expected, mean T-unit ratios are high all over (table 7.2). Far over 75% (that is: not exactly 'at least .80') of all words figure in flawless strings, be they correctly punctuated or not. Mean T-unit ratio in Description lags behind. The standard deviations of the mean T-unit ratios in Description pointedly hint at an enormous variance in students' performance. Contrariwise, mean T-unit ratios for Narrative are relatively high throughout, whereas standard deviations are relatively low. It seems as though Narrative as a genre, unlike Description, elicits flawless strings in a comparatively overwhelming amount.

1. So, in the experiment proper effects are computed over 750 words of each individual (average) student, in the extended experiment over 1250. For reference: Hunt (1965) collected 1000 words per student for his T-unit study.

Table 7.2. Experiment proper. T-unit ratio for Description, Narrative and Exposition separately and summed; means and standard deviation (s.d.)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
Description	n = 162		n = 164		n = 165	
- Direct Writing	.65	.36	.79	.23	.70	.31
- School Grammar	.70	.32	.73	.28	.75	.27
- Sentence Construction	.77	.29	.77	.28	.75	.29
Narrative	n = 159		n = 162		n = 162	
- Direct Writing	.86	.16	.86	.13	.87	.12
- School Grammar	.86	.17	.88	.13	.86	.13
- Sentence Construction	.83	.18	.87	.14	.88	.12
Exposition	n = 162		n = 164		n = 165	
- Direct Writing	.81	.24	.80	.23	.78	.21
- School Grammar	.73	.25	.73	.27	.75	.25
- Sentence Construction	.78	.24	.78	.24	.87	.16
Non-guided summed	n = 158		n = 162		n = 159	
- Direct Writing	.77	.16	.82	.13	.79	.13
- School Grammar	.76	.15	.78	.16	.78	.12
- Sentence Construction	.80	.14	.81	.15	.84	.12

As for the exploratory measures: mean length of flawless strings (table 7.3) hovers in between five and seven words. In narratives, mean T-unit length tends to be a trifle longer. Standard deviations are rather impressive throughout, with occasional outliers such as in Description at the pre test for Direct Writers and in Exposition at the mid test for School Grammarians.

Remarkably, though, when writing a narrative all students tend to produce T-units of more or less comparable length. Hardly any student scores zero (zero indicating that all strings were flawed) and none of them produce strings of extreme length: a flawless string of 13 words is the longest found in a narrative. For Description and Exposition the situation differs dramatically. In those genres there seems to be no average mean T-unit length: students differ enormously. Some produce no flawless strings at all. Five produced (fine) strings of extreme lengths such as 16 to 23 words in Description, five others did the same in Exposition. The other 130 had mean T-unit lengths anywhere in between.

Table 7.3.¹ Experiment proper. Mean T-unit length in Description, Narrative and Exposition separately and summed; means and standard deviation (s.d.)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
Description	n = 159		n = 162		n = 162	
- Direct Writing	5.63	5.00	5.99	3.05	5.02	2.63
- School Grammar	4.77	2.60	4.89	2.09	5.69	3.26
- Sentence Construction	5.54	3.05	5.44	3.55	5.65	3.54
Narrative	n = 161		n = 163		n = 162	
- Direct Writing	6.57	1.28	6.77	1.72	6.88	1.70
- School Grammar	6.44	1.76	6.70	1.58	6.79	1.73
- Sentence Construction	6.25	1.77	6.48	1.79	6.52	1.23
Exposition	n = 162		n = 163		n = 162	
- Direct Writing	6.08	2.86	6.25	2.24	6.21	2.22
- School Grammar	5.44	2.30	6.17	5.91	5.77	2.91
- Sentence Construction	6.00	2.81	6.63	3.74	6.31	2.81
Non-guided summed	n = 158		n = 163		n = 161	
- Direct Writing	6.09	2.08	6.33	1.52	6.03	1.54
- School Grammar	5.54	1.27	5.91	2.32	6.01	1.71
- Sentence Construction	5.99	1.63	6.17	2.16	6.19	1.86

Data on run-on ratios (table 7.4) tell almost the same tale as those for mean T-unit length. When writing a description or an exposition students differ enormously in their performance: some do not run-on at all, some run-on continuously, all others do something in between, none doing the same. When writing a narrative, though, performance of students is exactly the reverse: some students run-on to an extreme extent, most, however, run-on alike.

2. Number of subjects here or in any other table to come may not be equal to the number of subjects in Table 14. This is due to computational details, such as in this particular table, dividing zero by zero (in computing mean T-unit length of students that had not produced any words in syntactically well-formed strings).

Table 7.4. Experiment proper. Run-on ratio for Description, Narrative and Exposition, separately and summed; means and standard deviation (s.d.)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
Description	n = 159		n = 162		n = 162	
- Direct Writing	.21	.28	.22	.20	.26	.22
- School Grammar	.32	.40	.29	.25	.33	.28
- Sentence Construction	.32	.54	.24	.29	.24	.24
Narrative	n = 161		n = 163		n = 162	
- Direct Writing	.36	.21	.35	.17	.36	.20
- School Grammar	.36	.23	.38	.21	.43	.25
- Sentence Construction	.36	.21	.33	.23	.35	.20
Exposition	n = 162		n = 163		n = 162	
- Direct Writing	.26	.21	.27	.21	.24	.22
- School Grammar	.35	.33	.25	.24	.27	.22
- Sentence Construction	.25	.22	.22	.22	.20	.20
Non-guided summed	n = 158		n = 161		n = 161	
- Direct Writing	.28	.16	.27	.13	.29	.15
- School Grammar	.34	.25	.31	.14	.34	.15
- Sentence Construction	.32	.22	.26	.16	.26	.14

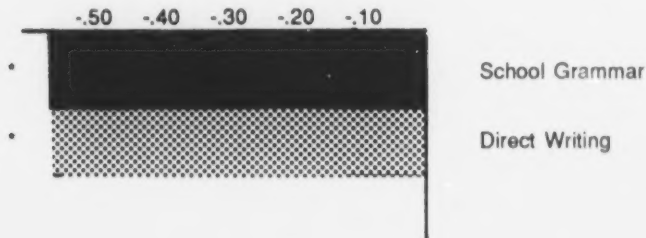
7.2 Academic linguistic knowledge

The Recognizing Verbs test assesses students' academic linguistic knowledge. Means and standard deviations on this test for the experiment proper are presented in table 7.5; figure 7.1 presents standardized contrasts.

Table 7.5. Experiment proper. Recognizing Verbs (discrete point test; n = 14); means and standard deviation (s.d.)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
	n = 161		n = 164		n = 165	
Direct Writing	3.89	1.63	4.69	2.05	5.20	2.04
School Grammar	3.22	1.22	4.06	1.65	5.18	1.54
Sentence Construction	3.37	1.72	4.47	2.43	6.05	2.11

Figure 7.1 Academic linguistic knowledge. Standardized contrasts for Recognizing Verbs



According to the second hypothesis put forward in this study, academic linguistic knowledge of Sentence Construction students equals that of School Grammar students and outstrips that of Direct Writing students. This hypothesis does not hold. As can be seen in figure 7.1, Sentence Construction students outperform the other students by slightly more than half a standard deviation on the test on Recognizing Verbs, whereas there is no effect at all of the other treatments. Recognizing verbs, by the way, seems to be very hard. At the close of the experiment proper students on average do not even answer half of the questions correctly.

Table 7.6.³ Extended experiment. Recognizing verbs (discrete point test; $m = 14$); means and standard deviation (s.d.)

	pre test		post test I		mid test II		post test II	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
	n = 96		n = 99		n = 97		n = 98	
writ-sent	3.91	1.70	5.13	1.88	4.55	2.40	6.25	2.71
sent-writ I	3.39	1.80	6.06	2.14	5.10	2.89	6.30	2.92
	n = 57		n = 60		n = 59		n = 59	
sent-writ II	3.61	1.77	6.17	2.18	5.31	2.65	6.59	2.73
gram-writ	3.55	1.45	5.06	1.44	3.53	2.06	5.50	2.43

Results of both follow-up studies (table 7.6) first of all discloses that ex-Direct Writing students catch up nicely with the ex-Sentence Construction students, whereas the latter do not seem to be plagued by amnesia as far as verbs are concerned. Next, it seems as though the writing ex-grammarians get the hang of

3. In tables, both follow-up studies will be referred to in shorthand throughout this chapter. Sentence Construction followed by Direct Writing and vice versa will be referred to as 'sent-writ' and vice versa; 'gram-writ' refers to School Grammar followed by Direct Writing. Also cf. table 6.3.

recognizing verbs, too. They appear to do so only after mid test II, that is to say, during the very last six or seven weeks of the extended experiment.

7.3 Sentence awareness

Data on sentence awareness are collected via two discrete point tests and in writing. The tests are Recognizing Sentences and Constructing Sentences. Awareness of sentences in writing is assessed by sentence ratio and explored mean sentence length. Tables 7.7 to 7.10 concern sentence awareness in the experiment proper, tables 7.11 to 7.14 concern the same in the extended experiment.

Table 7.7. Experiment proper. Recognizing Sentences (discrete point test; $m = 9$); means and standard deviation (s.d.)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
	n = 162		n = 164		n = 165	
Direct Writing	5.43	2.08	6.15	1.74	6.04	1.92
School Grammar	5.11	1.74	5.81	1.92	6.15	1.98
Sentence Construction	5.47	1.87	6.23	1.84	6.89	2.18

Table 7.8. Experiment proper. Constructing Sentences (discrete point test; $m = 10$); means and standard deviation (s.d.)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
	n = 159		n = 164		n = 165	
Direct Writing	5.21	2.42	7.30	2.25	7.80	2.07
School Grammar	5.15	2.30	7.43	2.21	7.98	1.70
Sentence Construction	5.31	2.46	7.86	2.34	7.43	2.39

Tables 7.7. and 7.8 bring out that sentence awareness in tests rapidly increases irrespective of treatment.

Students do not seem to encounter grave difficulties when they have to put genuine sentences to paper under guidance (table 7.9). However, in the three non-guided essays mean sentence ratio is low, no matter what genre students

write texts in. Far less than half of the words written down comes in student-wise demarcated genuine sentences.

Table 7.9. Experiment proper. Sentence ratio for Guided Description, Description, Narrative and Exposition, separately and summed over all three non-guided essays; means and standard deviations (s.d.)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
Guided Description	n = 162		n = 164		n = 165	
- Direct Writing	.78	.20	.83	.19	.78	.25
- School Grammar	.80	.19	.82	.16	.77	.20
- Sentence Construction	.85	.15	.86	.16	.89	.14
Description	n = 159		n = 161		n = 162	
- Direct Writing	.49	.39	.54	.37	.42	.37
- School Grammar	.46	.36	.45	.39	.43	.38
- Sentence Construction	.60	.34	.59	.36	.57	.34
Narrative	n = 161		n = 163		n = 163	
- Direct Writing	.44	.36	.40	.33	.38	.31
- School Grammar	.39	.33	.45	.35	.39	.32
- Sentence Construction	.47	.32	.50	.31	.55	.34
Exposition	n = 162		n = 163		n = 162	
- Direct Writing	.43	.36	.45	.37	.35	.30
- School Grammar	.39	.35	.42	.37	.42	.36
- Sentence Construction	.51	.37	.50	.38	.51	.36
Non-guided summed	n = 158		n = 162		n = 159	
- Direct Writing	.45	.27	.46	.28	.38	.26
- School Grammar	.41	.25	.44	.30	.40	.27
- Sentence Construction	.53	.25	.53	.26	.54	.28

Scores in sentence ratio seem to be a shaky bunch, means going up and down the way they do. Also note that none of the scores of students on sentence ratio seem to be connected in any way to their T-unit ratio scores. In other words, the amount of words segmented into sentences is in no way related to the amount of words in flawless strings. Instead of consulting appendices 2 and 3, where dependencies and twinsisterhood present a rather murky picture, one had better compare essays (5.1) and (5.2) to get an idea.

Finally, notice that Sentence Construction students seem to outperform the others already at the time of pre test, their overall score towering some .10 above scores of the other students. These high pre test scores of the Sentence Construction students are very conspicuous and may very well suggest these students definitely distinguish themselves from the others on one or the other very fundamental characteristic. They do not, however. These students are a random sample different in no way from the other two randomly sampled student groups. Their pre test scores are a mere freak of fortune.

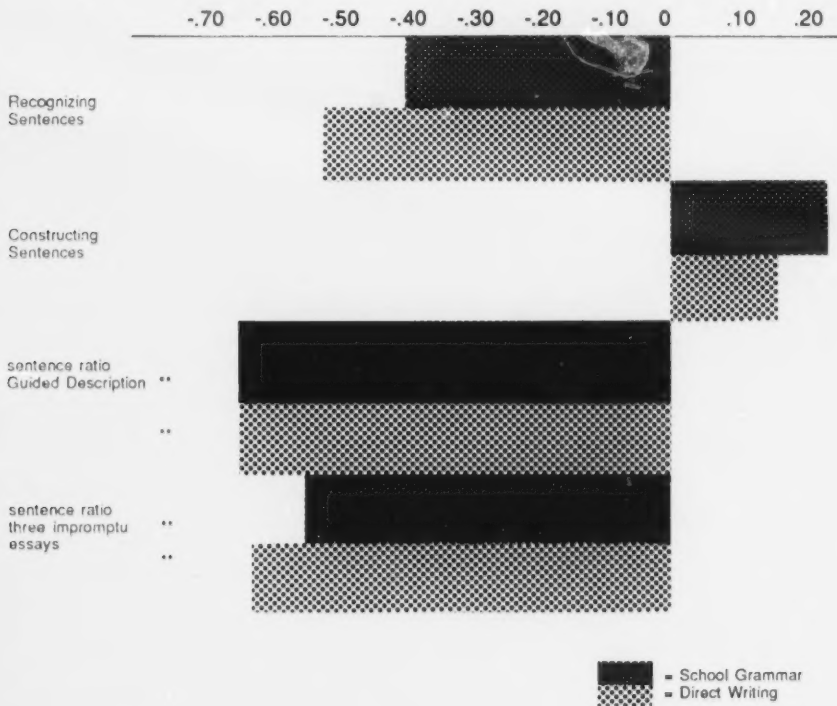
Table 7.10. Experiment proper. Mean sentence length for Guided Description, Description, Narrative and Exposition, separately and summed over all three non-guided essays; means and standard deviation (s.d.)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
Guided Description	n = 162		n = 164		n = 165	
- Direct Writing	5.33	.85	5.76	1.02	5.28	1.08
- School Grammar	5.25	.67	5.67	.95	4.81	.86
- Sentence Construction	5.63	1.12	5.59	1.15	5.13	1.03
Description	n = 159		n = 161		n = 162	
- Direct Writing	4.30	4.64	4.94	4.87	3.19	2.98
- School Grammar	3.27	2.80	3.39	2.68	4.05	4.18
- Sentence Construction	4.65	3.83	5.02	4.26	4.67	3.72
Narrative	n = 155		n = 152		n = 157	
- Direct Writing	4.98	3.08	5.07	3.28	5.09	3.74
- School Grammar	4.65	3.68	5.67	3.60	5.36	3.78
- Sentence Construction	5.43	3.29	5.60	3.38	5.91	3.22
Exposition	n = 162		n = 164		n = 164	
- Direct Writing	4.52	4.13	4.61	3.55	3.84	3.23
- School Grammar	4.56	5.60	5.10	6.69	5.00	4.93
- Sentence Construction	4.74	3.65	5.01	4.42	4.77	3.88
Non-guided summed	n = 152		n = 150		n = 154	
- Direct Writing	4.65	2.55	4.93	2.53	4.12	2.45
- School Grammar	4.13	2.53	4.89	3.09	4.74	2.84
- Sentence Construction	4.97	2.68	5.28	2.99	5.14	2.69

The measure exploring mean sentence length in impromptu essays found these lengths to amount to a meagre four to six words, means of course being pulled down mildly by students that write no or only one or two sentences without flaws. If these low scorers are not taken into consideration, mean length of sentences turns out to hover around some six to seven words. In Narrative and Guided Description mean length of sentences tends to be a trifle longer. In Guided Description test format and booklet are to be held responsible for the phenomenon, in Narrative students run on rather wildly, lengthening sentences by conjoining T-units (cf. table 7.4).

Figure 7.2 presents standardized contrasts of treatments on the scores gathered to measure sentence awareness. The upper half shows contrasts for the two discrete point tests, viz. Recognizing Sentences and Constructing Sentences, the lower half shows contrasts for sentence ratios of Guided Description and the summed sentence ratios for the three impromptu essays respectively.

Figure 7.2 Sentence awareness. Standardized contrasts for Recognizing Sentences, Constructing Sentences and the sentence ratios for Guided Description and the three impromptu essays



The hypothesis says that Sentence Construction students know what distinguishes sentences from non-sentences better than School Grammar or Direct Writing students do and, thus, will use more genuine sentences in their writing as well. The figure displays there are no effects of either of the three treatments on either test. This means the first half of the hypothesis does not hold whereas the second does: the figure brings out that in writing, whether guided or not, the Sentence Construction treatment achieves results that are far better than that of the other treatments, effects holding beyond the 99% percent interval.

Results of the extended experiment on sentence awareness are presented in tables 7.11 to 7.14. Tables 7.11 and 12 bring out that students, unaffected by treatment, work away happily on the tests on sentence awareness.

Table 7.11. Extended experiment. Recognizing Sentences (discrete point test; $n = 9$); means and standard deviation (s.d.)

	pre test		post test I		mid test II		post test II	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
	n = 97		n = 99		n = 98		n = 98	
writ-sent	5.47	2.10	6.00	1.96	7.56	1.62	6.58	1.52
sent-writ I	5.52	1.91	6.86	2.22	7.78	1.79	7.38	1.47
	n = 58		n = 60		n = 59		n = 59	
sent-writ II	5.69	1.75	7.07	2.02	7.93	1.56	7.55	.99
gram-writ	5.78	1.59	6.43	1.87	7.37	1.87	6.93	1.44

Table 7.12. Extended experiment. Constructing Sentences (discrete point test; $n = 10$); means and standard deviation (s.d.)

	pre test		post test I		mid test II		post test II	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
	n = 96		n = 99		n = 96		n = 96	
writ-sent	5.17	2.15	7.75	2.08	8.08	1.57	8.56	1.76
sent-writ I	5.47	2.47	7.43	2.48	8.22	2.22	8.66	1.62
	n = 55		n = 60		n = 59		n = 59	
sent-writ II	5.71	2.73	7.43	2.14	8.31	1.87	8.79	1.37
gram-writ	5.74	2.18	7.93	1.70	8.27	1.70	8.17	1.66

Tables 7.13 and 7.14 present means and standard deviations for sentence ratio, the main measure for students' awareness of sentences in writing, and mean sentence length, its exploratory partner.

The latter table brings out mean sentence length remains unaffected. The former, 7.13, brings out Sentence Construction does not seem to be able to influence sentence ratios of ex-Direct Writers in Guided Description as it did the sentence ratios of the Sentence Construction students proper. What is more, its beneficial effects seem to drain away when treatment is withheld. The latter phenomenon is manifest in the impromptu essays as well (consult the lowest row of data, on the summed means and standard deviations of non-guided writing), though here the Sentence Construction programme seems to succeed in modestly propelling sentence ratios of ex-Direct Writers. There is no indication whatsoever, then, that the hypothesis for this first follow-up study holds: students' sentence awareness, particularly in writing, does not at all thrive most if they are taught to construct sentences first and to construct text next.

It seems as though, indeed, School Grammar followed by Direct Writing comes to nothing as far as sentence awareness goes, as the hypothesis for the second follow-up study boldly states. But then, it would be rather nonsensical to

maintain that Sentence Construction followed by Direct Writing works wonders. Withholding Sentence Construction instruction reduces students to their initial incapability, or so it seems.

Table 7.13. Extended experiment. Sentence ratio for Guided Description, Description, Narrative and Exposition, separately and summed over all three non-guided essays; means and standard deviation (s.d.)

	pre test		post test I		mid test II		post test II	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
Guided Description								
writ-sent	n = 97	.78	n = 99	.77	n = 98	.76	n = 97	.78
sent-writ I	.20	.85	.23	.89	.26	.88	.26	.85
		.16		.15		.15		.18
sent-writ II	n = 58	.86	n = 60	.90	n = 59	.87	n = 59	.82
gram-writ	.15	.82	.13	.73	.16	.75	.20	.75
		.18		.25		.23		.21
Description								
writ-sent	n = 96	.51	n = 98	.41	n = 98	.44	n = 99	.50
sent-writ I	.38	.62	.37	.58	.39	.51	.35	.58
		.34		.35		.36		.35
sent-writ II	n = 55	.54	n = 59	.55	n = 59	.50	n = 60	.53
gram-writ	.36	.54	.35	.36	.34	.27	.37	.47
		.35		.35		.34		.37
Narrative								
writ-sent	n = 96	.46	n = 97	.38	n = 97	.38	n = 98	.45
sent-writ I	.35	.48	.32	.56	.34	.41	.36	.37
		.32		.34		.32		.36
sent-writ II	n = 57	.46	n = 58	.54	n = 59	.41	n = 60	.33
gram-writ	.32	.44	.34	.37	.32	.27	.32	.28
		.35		.35		.34		.32
Exposition								
writ-sent	n = 97	.41	n = 97	.35	n = 98	.46	n = 98	.50
sent-writ I	.36	.52	.30	.52	.37	.43	.34	.48
		.37		.35		.36		.37
sent-writ II	n = 58	.45	n = 59	.48	n = 59	.43	n = 60	.42
gram-writ	.35	.48	.36	.40	.37	.27	.35	.38
		.33		.39		.38		.35
Non-guided summed								
writ-sent	n = 95	.46	n = 96	.38	n = 97	.43	n = 97	.48
sent-writ I	.27	.54	.27	.56	.29	.45	.29	.47
		.25		.27		.26		.26
sent-writ II	n = 54	.50	n = 56	.52	n = 59	.45	n = 60	.43
gram-writ	.25	.48	.27	.36	.25	.28	.26	.38
		.26		.25		.28		.26

Table 7.14. Extended experiment. Mean sentence length for Guided Description, Description, Narrative and Exposition, separately and summed over all three non-guided essays; means and standard deviation (s.d.)

	pre test		post test I		mid test II		post test II	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
Guided Description								
writ-sent	n = 97		n = 99		n = 97		n = 97	
sent-writ I	5.30	.86	5.15	1.02	4.94	1.19	5.25	1.09
	5.58	1.12	5.07	.99	5.31	.89	5.32	.98
sent-writ II	n = 58		n = 60		n = 59		n = 59	
gram-writ	5.57	1.03	5.25	.99	5.33	.78	5.26	.94
	5.26	.62	4.73	.79	5.07	.88	5.68	1.76
Description								
writ-sent	n = 96		n = 97		n = 97		n = 98	
sent-writ I	4.39	4.64	3.09	2.95	3.69	3.30	3.63	2.50
	4.72	3.87	4.67	3.69	4.19	3.24	5.13	4.96
sent-writ II	n = 55		n = 58		n = 58		n = 59	
gram-writ	4.26	4.16	4.09	3.03	3.70	2.75	4.84	5.79
	3.45	2.54	3.39	3.52	3.22	3.98	4.34	3.24
Narrative								
writ-sent	n = 93		n = 95		n = 90		n = 95	
sent-writ I	5.19	2.98	4.82	3.65	5.25	3.66	5.28	3.84
	5.19	2.98	5.75	3.20	4.93	3.25	4.71	3.87
sent-writ II	n = 54		n = 57		n = 56		n = 56	
gram-writ	5.25	2.91	5.36	3.21	4.70	3.18	4.33	3.14
	4.94	3.40	5.08	4.08	3.62	4.05	3.90	3.43
Exposition								
writ-sent	n = 97		n = 97		n = 98		n = 98	
sent-writ I	3.99	3.25	3.77	3.23	4.78	3.92	5.73	5.84
	4.65	3.53	4.86	3.90	4.36	3.24	5.46	4.94
sent-writ II	n = 58		n = 59		n = 58		n = 60	
gram-writ	4.37	3.07	3.96	3.09	4.07	2.80	4.75	4.33
	6.27	6.75	4.54	4.97	2.48	3.37	4.29	3.54
Non-guided summed								
writ-sent	n = 92		n = 93		n = 89		n = 93	
sent-writ I	4.57	2.41	3.98	2.42	4.68	2.84	4.92	2.86
	4.87	2.50	5.05	2.69	4.65	2.34	5.26	3.15
sent-writ II	n = 51		n = 54		n = 54		n = 55	
gram-writ	4.63	2.23	4.36	2.12	4.38	1.84	4.83	3.03
	4.91	2.76	4.31	2.66	3.22	3.16	4.22	2.32

7.4. Conciseness

Data on students' conciseness were collected via one test (Informing Succinctly) and in writing (noun phrase complexity ratio). Means and standard deviations are presented in tables 7.15 and 7.16.

Table 7.15. Experiment proper. Informing Succinctly (discrete point test; $n = 10$); means and standard deviation (s.d.)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
	n = 160		n = 164		n = 166	
Direct Writing	2.53	2.19	3.31	2.63	4.93	3.26
School Grammar	1.85	2.04	3.13	3.33	4.29	3.41
Sentence Construction	2.57	2.48	4.77	3.34	5.86	3.50

Table 7.16. Experiment proper. Noun phrase complexity ratio for Guided Description, Description, Narrative and Exposition, separately and summed over all three non-guided essays; means and standard deviation (s.d.)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
Guided Description	n = 157		n = 164		n = 165	
- Direct Writing	.25	.17	.30	.18	.28	.19
- School Grammar	.27	.15	.30	.16	.21	.17
- Sentence Constr.	.26	.15	.34	.18	.34	.21
Description	n = 159		n = 162		n = 163	
- Direct Writing	.36	.25	.26	.24	.28	.21
- School Grammar	.26	.19	.31	.21	.36	.25
- Sentence Constr.	.26	.22	.26	.23	.26	.25
Narrative	n = 161		n = 163		n = 164	
- Direct Writing	.19	.13	.17	.12	.16	.12
- School Grammar	.15	.11	.17	.13	.20	.16
- Sentence Constr.	.18	.14	.16	.13	.17	.11
Exposition	n = 161		n = 163		n = 162	
- Direct Writing	.22	.19	.22	.16	.21	.18
- School Grammar	.16	.17	.21	.21	.14	.14
- Sentence Constr.	.17	.17	.20	.17	.20	.21
Non-guided summed	n = 157		n = 162		n = 161	
- Direct Writing	.26	.13	.22	.09	.21	.10
- School Grammar	.19	.09	.23	.11	.23	.12
- Sentence Constr.	.20	.11	.20	.12	.21	.12

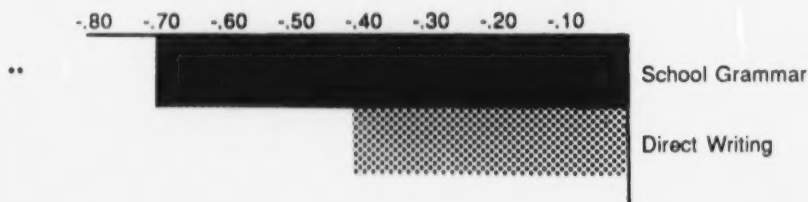
The very low means for the test (table 7.15), even at the third time of test taking, point out that students encounter grave difficulties when they have to inform succinctly.

As constructing complex noun phrases very emphatically is not their favourite pastime, students do not spontaneously generate herds of these in their writing, as table 7.16 testifies. There seems to be a marked difference between both descriptions and the other two assignments: both means and standard deviations are higher when students write a description.

The hypothesis on conciseness has it that Sentence Construction students are attuned to squash information in noun phrases more readily than School Grammar and Direct Writing students. So, it goes on, Sentence Construction students are expected to employ complexly structured noun phrases in their writing more frequently than School Grammar and Direct Writing students do. The psychometric reliability of the complex noun phrases ratio, in all probability due to the utter lack of variance, forbids a proper test on the latter part of the hypothesis. Means and standard deviations, though, point out that in all likelihood the hypothesis is wildly off the mark.

The test does allow computing standardized contrasts. These are presented in figure 7.3. This figure shows School Grammar students to perform worse than Sentence Construction students. The difference holds beyond the 99% interval. There is no difference between Direct Writing students and their Sentence Construction colleagues: effects of both treatments amount to the same. Evidently, two thirds of the hypothesis on conciseness was wrong.

Figure 7.3 Conciseness. Standardized contrasts for Informing Succinctly



Psychometric unreliability renders presenting scores on noun phrase complexity ratios in the extended experiment useless. So, scores on the Informing Succinctly test only are presented.

Table 7.17 brings out all the more that Direct Writing beneficially influences conciseness of students. The ex-School Grammar students entered the follow-up experiment well behind the other students. Apparently Direct Writing treatment enhances their test performance. Ex-School Grammarians do not, however, catch up completely. So, as Sentence Construction quite unexpectedly finds its match

in Direct Writing when it comes to affect conciseness, hypotheses on both follow-up studies had best be relegated to the realm of fantasy.

Table 7.17. Extended experiment. Informing Succinctly (discrete point test; $n = 10$); means and standard deviation (s.d.)

	pre test		post test I		mid test II		post test II	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
	n = 96		n = 99		n = 98		n = 98	
writ-sent	2.47	2.26	4.73	3.31	6.35	3.20	8.08	2.61
sent-writ I	2.27	2.32	5.67	3.53	6.68	3.05	7.88	2.53
	n = 57		n = 60		n = 59		n = 59	
sent-writ II	2.51	2.50	5.83	3.64	6.79	3.08	8.17	2.19
gram-writ	1.85	1.88	3.83	3.04	5.60	3.15	7.00	3.27

7.5 Overall writing quality

The final hypothesis of this study holds that Sentence Construction students write impromptu essays superior in overall quality to those of Direct Writing students and that these in turn are superior in overall quality to those of School Grammar students. Means and standard deviations, presented in table 7.18, however, suggest that overall writing quality of students' prose is relatively autonomous and not liable to treatment or to substantial changes over short periods of time.

Table 7.18. Experiment proper. Overall quality scores for Description, Narrative and Exposition; (lowest score: 1; highest: 5); means and standard deviation (s.d.)

	pre test		mid test I		post test I	
	mean	s.d.	mean	s.d.	mean	s.d.
Description	n = 159		n = 164		n = 163	
- Direct Writing	2.90	.79	3.00	.84	2.92	1.00
- School Grammar	2.87	.73	3.08	.91	3.03	.90
- Sentence Construction	3.09	.83	2.98	.95	2.91	.96
Narrative	n = 162		n = 164		n = 165	
- Direct Writing	2.76	.91	3.10	1.01	2.82	.96
- School Grammar	2.96	.92	2.97	.94	2.62	.86
- Sentence Construction	2.74	1.00	2.81	.98	2.89	1.02
Exposition	n = 162		n = 164		n = 161	
- Direct Writing	2.88	.80	2.78	.93	2.79	.87
- School Grammar	2.70	.86	2.71	1.00	2.75	.99
- Sentence Construction	2.72	1.00	2.86	1.02	2.92	1.05

The table suggests quality of all genres was equal. Such a conclusion is not warranted, though, as ratings of the three genres were performed separately, rendering comparison of quality between genres impossible. If, for that matter, such a comparison is possible at all⁴.

Standardized contrasts for the last of the five variables are presented in figure 7.4. This is a miniature figure, two dwarfish flags waving left of the pole, indicating that, whatever the effects of whatever treatment, overall quality remained unchanged for all students.

Figure 7.4 Overall writing quality. Standardized contrasts summed over three impromptu essays

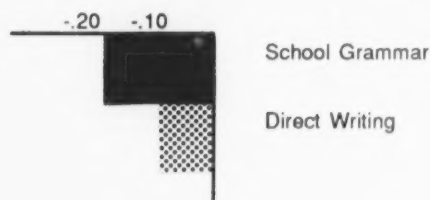


Table 7.19 reports on means and standard deviations of the overall quality scores for the students in the extended experiment. The table bears no evidence to the felicitousness of whatever ordering of treatments. So, once again, it seems hypotheses concerning the extended experiment are a wee bit overstrung. The Sentence Construction course does not turn out to provide the fecund soil it was said to provide, let alone the course works wonders.

7.6 Summary

This chapter presented results of both experiments engaged in this study. The experiment proper involved 165 subjects of 6 schools. Per school, subjects were assigned randomly to one of three treatments, viz. to either a course in Direct Writing, a course in School Grammar or one on Sentence Construction. The extended experiment comprised two modest studies into follow-up effects. The first of these involved 99 students, 48 of which first learned to construct sentences and next were taught to construct texts. The other 51 first learned the latter and were taught to construct sentences only after that. The second study involved sixty students, half of which were familiarized with School Grammar

4. And this is exactly why overall quality of texts was rated separately per genre.

Table 7.19. Extended experiment. Overall quality scores for Description, Narrative and Exposition; means and standard deviation (s.d.)

	pre test		post test I		mid test II		post test II	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
Description	n = 97		n = 98		n = 98		n = 98	
writ-sent	2.93	.80	3.02	.99	2.83	.85	2.73	.92
sent-writ I	3.10	.86	2.83	.96	2.86	.86	2.83	.94
	n = 55		n = 59		n = 59		n = 56	
sent-writ II	3.18	.80	2.82	1.03	3.07	.87	2.91	.93
gram-writ	2.94	.78	2.92	.92	2.90	.79	2.90	.83
Narrative	n = 97		n = 99		n = 98		n = 98	
writ-sent	2.78	.95	2.83	.95	2.82	1.04	2.73	1.11
sent-writ I	2.65	.96	2.89	1.03	3.06	1.01	2.66	1.23
	n = 57		n = 60		n = 59		n = 59	
sent-writ II	2.84	.98	2.75	1.04	3.20	.73	2.86	1.12
gram-writ	3.07	1.02	2.67	.82	2.69	.82	2.68	.88
Exposition	n = 97		n = 98		n = 97		n = 97	
writ-sent	2.85	.80	2.74	.90	2.80	.81	2.65	.79
sent-writ I	2.72	1.01	2.89	1.06	2.64	.97	2.58	1.03
	n = 58		n = 60		n = 58		n = 60	
sent-writ II	2.77	1.13	2.70	1.12	2.49	1.06	2.65	1.14
gram-writ	2.95	.97	2.55	.97	2.68	.94	2.68	.85

first, while the other half qualified themselves in Sentence Construction. For a follow-up, all were trained in writing via the Direct Writing course. The experiment proper lasted thirteen weeks; classes were taught twice a week and took one hour each. As the extended experiment comprises both the experiment proper and a follow-up treatment, it took exactly twice as long.

Students' writing fluency was measured by analyses of T-unit ratio. It was supposed to be fine and no treatment was supposed to affect it. It was fine, indeed, and, it seems, no treatment affected it. There were minor effects of genre, highlighted in results of Narrative as opposed to Description. Narrative as a rule resulted in the highest T-unit ratios, Description in the lowest. The narrative genre elicited most text and longest T-units opposed to, again, the description, the latter on the average eliciting little text in rather short strings. Students' academic linguistic knowledge was assessed by one test on Recognizing Verbs. Contrary to hypothesis, students' academic linguistic knowledge profited from the course on Sentence Construction only, both in the experiment proper and in the extended experiment. Information on sentence awareness is gathered via discrete point tests on the one hand and via writing on the other. Results revealed that the first half of the

hypothesis on sentence awareness does not hold: Sentence Construction students do not outperform the other students. All students are equally aware of what are sentences and what aren't. In writing the course on Sentence Construction decidedly enhances students' sentence awareness. That means the second half of the hypothesis holds: Sentence Construction students, provided they are actually working with the course, outperform others.

Sentence Construction instruction boosted students' conciseness. So did the writing course. So here is another erroneous hypothesis. Effects of treatment on conciseness in writing could not be computed.

The final hypothesis, on the beneficial influence of Sentence Construction and the slightly less but all the same beneficent effect of Direct Writing, was faulty, too. Quality of the essays did not profit at all.

DISCUSSION

... children are not little linguists ... because they use a cognitive system different from the one that is commonly used by the linguist ... This assumption also seems to explain why the child is consistently successful, while the linguist is not ...

Felix (1987)

This chapter discusses the horde of results, presented in chapter 7, of the experiment designed to study the effects of grammar instruction upon junior writing. In the experiment, students were assigned randomly to conditions, so there was no contamination of school and treatment. Neither were conditions contaminated with experimenters: all three experimenters taught all courses. Such a randomized block design is a first safeguard in experimenting, ruling out school and experimenter effects. Yet in practice other dangers are lurking, still. Experimenters may follow their own lead and not that of the course, students may spend far too much time on matters not at all relevant to the course or the course may be miles above their heads. Nuisance variables such as these were controlled for and chapter 6 brought out that protocol adherence of experimenters was satisfactory, that all students except perhaps the Sentence Construction students in the second experiment spent enough time on course material, that students in both the School Grammar and the Sentence Construction treatment easily met the set mastery norm and that there was no systematic interference of the regular language arts programs. Moreover, there were no systematic differences between conditions in students' reading skill and, finally, there were no experimenter effects. So, whatever the effects will turn out to be, they certainly are experimental effects and nothing else.

Data have been analyzed by computing standardized contrasts over scores gathered by psychometrically reliable tests and measures only. The intervals that go with these contrasts are another safeguard, indicating the degree of confidence to be put in results. So, this chapter will engage in a detailed discussion of results knowing that it is quite safe to take all of these for what they should be: results of experimental treatment.

The experiment meant to test a theory on grammar and writing. In its most general terms this theory says that grammar instruction may very well be a

means to boost student writing. That is: if it is the kind of knowledge students want while writing. Such knowledge is, e.g., knowledge on sentences. As the fully-fledged sentences of written language do not come naturally, language users ought to be instructed explicitly on what are and what aren't sentences.

In order to test this theory, instruction on sentences in the study came in an all-grammatical, a-rhetorical format. It was hypothesized that, though taught outside the writing curriculum, a course on sentence construction, highlighting sentence patterns as a means to structure content, would bring about sentence awareness and readily available while writing at that. This, in its turn, was supposed to ameliorate conciseness and quality of writing.

This study also meant to demonstrate that the sort of grammar traditionally associated with grammar instruction in school may teach students to parse sentences and recognize parts of speech but is useless in writing.

Also, the study meant to demonstrate that a mere course in writing, whatever its merits be, will not help students to get to grips with the language writing comes in. It might help writers to search and organize content, it might help them to ponder over goal and public, it does, however, not familiarize them with the linguistic peculiarities of written language, a foreign language of a kind.

Now that results are known, what of this theory then? Did the study demonstrate that grammar instruction is indeed a means to boost student writing? It certainly did, though this should be put a trifle more circumspectly: though the Sentence Construction course did not ameliorate conciseness and quality of impromptu essays, it successfully triggered readily available awareness of genuine sentences in writing.

And did it demonstrate the utter uselessness of teaching traditional grammar as a means to ameliorate writing? It decidedly did that, too, and more to boot: it demonstrated that the School Grammar course did not even result in any academic linguistic knowledge to speak of.

The study also brought out that, according to theory, the 'mere' Direct Writing course did not familiarize novice writers the linguistic peculiarities of written language.

There is, however, far more than meets the eye in this one fell swoop. Remember that results of tests on sentence awareness suggested that sentences are not as unnatural as all that: the trio of experimental groups all scored equally successfully. What is amiss, here?

Also it turned out that the School Grammar treatment did not do the slightest bit of good to students' academic linguistic knowledge. This surely is a puzzle that needs to be solved. Moreover, chapter 7 brought to light that the Direct Writing course did not affect junior writing positively though it turned out that the course equalled the course on Sentence Construction when it came to get junior writers to inform succinctly. How did it do that?

And what is it that ails writing fluency of juniors, otherwise fine, when it comes to write a description?

This chapter, then, engages in a lengthy discussion of matters such as these. All five variables will pass by in Indian file once again and, in an attempt to penetrate to the heart of the matter, each will be shot at and held up to the light one last time. So, this chapter mainly discusses validity of instruments for dependent and independent variables, as some would rather have it. Be that as it may, by turning data inside out, it will bring to the fore essential flaws, mild surprises and a host of new questions. And some answers, too.

8.1 Inviolable writing fluency

As writing fluency originates in the (tacit) knowledge the writer has of his mother tongue it was hypothesized to be fine. No treatment was supposed to affect it. Novices' writing fluency was assessed by computing T-unit ratio, i.e. the ratio of words that come in flawless clauses and all words in the essay. This T-unit ratio was meant to sketch some sort of a baseline, corroborating the hypothesis that young writers encounter no difficulties in writing correct syntactic structures. Underlying this hypothesis is the assumption that all apprentices are old hands at grammar and all equally skilled at writing fluently. This does not mean all writing is flawless, though. Attention may occasionally falter and this may cause words to appear twice, or words to appear not at all, or structures to blend. Knowledge may fall short and this may result in deviantly structured strings. In an assessment of writing fluency these errors, whether accidental or resulting from lack of knowledge, should not be taken to be irrelevant. After all, they are there, whereas they should not. So, texts, or, to be more precise, the strings the text was made of, were analyzed the way they came in writing. This is contrary to tradition, that professes there are correct T-units only. This study knew correct as well as incorrect T-units; slightly maimed strings did not get repaired, seriously bruised ones were not done away with.

Six kinds of errors might trouble strings (see chapter 4.2 and 4.4). As can be seen from table 8.1, the one most frequently encountered was omission of a word or two from a string. Exposition is plagued by other errors most: conversational elements intrude this kind of prose relatively frequently. In descriptive texts omissions are especially prominent. Remember that chapter 7.1 already pointed out that mean T-unit ratio in Description 'lagged behind', implying that this genre was characterized by faulty strings. Table 8.1 brings out that in the descriptions students wrote a rather overwhelming amount of words is missing. Why should especially descriptions be troubled thus?

Table 8.1. Troubles in writing fluency per genre in percentages* (DES = Description; NAR = Narrative; EXP = Exposition)

	DES	NAR	EXP
total amount of errors	22	11	19
strings missing a word (or two)	18	5	11

Also, in chapter 7.1, standard deviations of the mean T-unit ratios in Description were said to hint pointedly at the enormous variance in students' performance as opposed to the narrative genre in particular. What causes this enormous variance in Description? The simplest way to gain insight into both matters is to take a look at the descriptions students wrote. Here are three. Texts (8.1) and (8.2) both describe one of the children's portraits (figure 2.2); (8.3) describes one of the mugs (figure 2.1). Text (8.1) scores 1.00 on T-unit ratio. This implies all strings are flawless. Text (8.2) scores .63. This means that only slightly over half of its strings are flawless. The description of the mug, (8.3), comes in defective strings only. Its T-unit ratio is zero.

- (8.1) (T-unit ratio: 1.00)
 the boy has a crew cut
 and he has got a small bear in his hands
 and he wears a sweater that is striped
 (1/23/18/M1; Kamal)
- de jongen heeft stekels
 en hij heeft een beertje in zijn handen
 en hij heeft een trui aan die strepen heeft
- (8.2) (T-unit ratio: 0.63)
 it is a boy
 curly hair
 two ears
 two eyes
 one nose
 one mouth
 he has two sacks
 and two stripes at the sides and one
 stripe mid
 he looks happy
 and he has two buttons above his
 sacks
 and I see the neck
 (1/12/02/M1; Yilmaz)
- het is een jonge
 krulle haar
 twee oren
 twee ogen
 één neus²
 één mond
 hij heb twee zakken
 en twee streep aan de zijkant èen
 streep mide
 hij kijkt blij
 en hij heb twee knopen boven zijn
 zakke
 en ik zie de nek
- (8.3) (T-unit ratio: 0.00)
 Three clouds and four trees.
 Three small houses a lot of
 small houses together.
 And a locomotive and three wagons
 Three sheep.
 (3/13/04/M1; Geeske)
- Drie wolken en vier bomen.
 Drie huisjes een heleboel
 huisjes aan elkaar.
 En een locomotief en drie wagonen
 Drie schapen.

1. Results are given irrespective of test period or condition, as neither affect T-unit ratio.

2. The acute accent does not add stress but distinguishes the Dutch article *een* ('a') from the numeral *een* (*one*).

The three texts, mere lists of elements as descriptions typically are, are adequate descriptions, and none too bad at that. Yilmaz' description is one of almost exact average overall quality: mean equals 2.98, her score was 3.00. The other two are one standard deviation below mean, Kamal scoring 2.00, Geeske scoring 2.24.

The most prominent characteristic distinguishing (8.1), flawless, from the completely faulty (8.3) is that (8.3) is essentially an unstructured list, consisting of merely threaded elements, framed in noun phrases only, whereas (8.1) lists elements by stringing full clauses or S-es. Add two words to all 'sentences' in (8.3) (I see; or: There are) and Geeske's T-unit ratio would be saved. Yilmaz' description, (8.2), employs both full clauses and noun phrases; her T-unit ratio score tumbles down proportionally, landing her somewhere in the middle.

This straightforwardly explains what ails the genre: as describing is essentially listing, students list. This, evidently, is simply an effect of the genre. This genre effect neatly explains the enormous variance as well: both Kamal, author of (8.1), and Geeske, author of (8.3) may be considered to be exponents of main streams in student writing. Though the listers decidedly are a minority, they are always there'. By steadily scoring zero on T-unit ratio they very effectively pull down its mean and gear up its variance.

It may seem a bit of a pity that T-unit ratio classifies lists, themselves consisting of perfectly fine noun phrases, as erroneous strings. Erroneous, that is, in the same sense that strings missing an in all probability in haste forgotten word or two are erroneous. After all, omitting all other phrases in order to merely make a list or omitting a word due to a lapse in attention may very well be argued to be very different errors essentially. Perhaps this had best be illustrated by considering what comes in when either of the two omissions have to be restored. Restoring an accident poses no problem to writers: by keenly rereading their prose they will detect the shortcomings and iron them out without much ado. Changing a list of noun phrases into full clauses, however, will require consciously controlled manipulation of syntactic structures of a kind. But far more important seems to be the writer's realization he needs to so change his writing. Such a realization implies knowing that writing is, in some respects, a language with rules that differ from that of speech.

However, as chapter 5.2 argued, such a fine-grained distinction of errors was not what the T-unit ratio in this study was meant for. T-unit ratio should simply separate flawless full clauses from the flawed ones in order to display what is and is not a problem for apprentice writers. And that is what it does rather adequately: it convincingly shows that novices produce full clauses pretty fluently.

3. Which does not mean they are the same students always. They are not.

8.2 Kernels vs. universals: academic linguistic knowledge

Academic linguistic knowledge is the kind of knowledge traditionally associated with grammar instruction in school: knowledge on parsing sentences and recognizing parts of speech. It is declarative and taught explicitly. This study hypothesized that School Grammar students and Sentence Construction students would hold the same amount of it. Academic linguistic knowledge of the School Grammar kind would, however, prove to be an inert body of knowledge as far as writing goes and, thus, the utmost the course would result in at that, whereas their familiarity with academic linguistic knowledge would mean Sentence Construction students were provided with the equipment and language to evaluate sentences in writing. As a matter of course, the Direct Writing programme would result in no such knowledge at all.

Chapter 7.2 brought out that at the close of the experiment proper Sentence Construction students outperform both other groups of students by slightly more than half a standard deviation on the test on Recognizing Verbs. Neither of the other two treatments yield results. The Direct Writing course was never supposed to. The School Grammar course was, though. How come its effects fail to appear?

The design of the experiment rules out school and experimenter effects as causes for this. Neither were all low-achievers accidentally assigned to the School Grammar treatment. What is more, chapter 6.2 brought out that, like the other students, School Grammar students decidedly got all the instruction they were meant to get in the way they were meant to get it. Also, they took it in as they should have. Finally, no instruction interfered. So, as there is neither anything the matter with either the psychometric reliability of the test or the analysis run over its scores, there is only one explanation left: the programme itself is responsible. Where does it fail, then? What was it that students took in so pernickety?

The School Grammar course, it says in chapter 4.3, focuses on parsing sentences and on recognizing parts of speech. It introduces traditional terminology late or not at all and freely uses terms coined on whatever bases that turn out convenient. Verbs are a case in point, as the course, avoiding the term 'verb', coined these 'now or yesterday words' or words that tell 'when', thus, says chapter 4.3, 'hammering home the semantics of the finite verb while in the mean time cramming students for one of the crucial heuristics in identifying this verb form'. As one of the major aims of the course is having students recognize the 'kernel sentence' (Du: de zinskern), identifying this verb form is of no small importance in the course. Now, notice that not verbs as a category of words, but rather finite verbs, or the way these particular forms of a part of speech function in a sentence, are a key issue, then. The introduction of the term 'finite form' (Du: persoonsvorm), preceding the term 'verb', corroborates this conclusion. Also notice that the heuristic for identifying the word seems to be attuned very delicately to this sole function in the sentence. After all, identifying words that tell 'when' is impossible when these are presented out of the

context where they actually tell 'when'. Also, infinitives or participles do not tell 'when'; neither are these forms introduced to students. These students, then, are familiar with one of the specialties of the verb in a sentence, viz. tense, but not with the hallmark that singles out verbs as a major category, viz. their being the words that name what people, animals or things do. So, by evaluating only a part of its contribution to the meaning of the sentence, these students are perfectly well able to recognize a 'finite form' in a sentence. As, however, they do not know what it is a finite form of; as, in other words, they are familiar only with part of the semantics of the verb, they are unable to do what the test (cf. chapter 5.3.) wants them to: distinguish verbs as a major category. They simply do not know what verbs are. And indeed, this is precisely where School Grammar students fail.

A burning question, then, is whether this School Grammar course is a valid course in school grammar. And the ardent answer to this question is yes, it is. As chapter 4.3. has already argued, the course is the best representant of the regular courses on traditional school grammar in a modern outfit. It is 'modern', in that it avoids crushing students with tons of terms, and it is traditional in that it familiarizes students with tons of notions all the same. It is modern in that it straightforwardly works towards recognition of the 'kernel sentence'. It is traditional in that it professes that such academic linguistic knowledge will beneficially influence language usage of students. It is more than a mere average of what students are forced to swallow in Dutch school grammar courses as it is based on *Jouw Taal Mijn Taal*. The grammar component of this course is head and shoulders above those of other courses.

Notice how the heuristic for verb recognition in the School Grammar course contrasts sharply to the heuristic for verb recognition in the course on Sentence Construction. Here, too, the function of the verb in the sentence is a key issue. What is more, verbs are what sentences are actually patterned on. But then, students start to learn about this only after they have familiarized themselves with what it is that distinguishes verbs from other parts of speech. Next, identifying the verb in a sentence crucially involves thinking over what it is the sentence tells, and, if the sentence tells what people do, what it is that these people do, exactly. Tense is not singled out in any way but is touched upon as a matter of course. After all, there are sentences that tell what people did, too. The two courses are contrasted likewise in the way they deal with nouns and adjectives. In the School Grammar course, nouns are words that go with an article (Du: *de, het* or *een*), adjectives are words that 'say something about the noun' (Du: *het zegt iets over het zelfstandig naamwoord*). In the Sentence Construction course nouns are words that name people, animals or things, adjectives are words that name what you think of them or what they look like. The Sentence Construction course, then, seems to bring to the fore the universal characteristics distinguishing major categories: actions (verbs) versus entities (nouns) versus qualities (adjectives). In the test on Recognizing Verbs distinguishing major categories is exactly what students need to do. The test has

them tick the one word of four (viz. three lexical verbs and one noun or one adjective) that is 'different' and 'should not be there' (also cf. chapter 5.3). The test brings out that the Sentence Construction students evidently are the only ones that got the hang of distinguishing major categories.

Results of the extended experiment bear evidence to this conclusion. Table 7.6 brought out that the ex-Sentence Construction students enrolled in the Direct Writing course in the second experiment kept on doing fine on the test, whereas the students treated the other way round quickly caught up with them. Remember, however, the table also showed that performance of the ex-School Grammarians enrolled in the course on Direct Writing slowly, slowly improved. It did so after mid test II only, that is to say, during the very last six or seven weeks of the second experiment. If, as the results bring out, neither School Grammar nor Direct Writing teach students what are parts of speech, then this is a very mysterious phenomenon indeed. After all, the Sentence Construction course was the only course properly guiding students towards a correct distinction of the parts of speech. Students not so guided (i.e. these writing ex-grammarians) need to make the proper distinctions themselves. And they seem to succeed in doing so at last.

This may mean the School Grammar course did do some good to students' academic linguistic knowledge after all. It took a while, the effect is small, but here it is. It may also mean that the School Grammar course did do no good to students' academic linguistic knowledge. Students may very well have succeeded in making the proper distinctions all by themselves. Their taking the test for the fifth time running might even be responsible for this small success. The test admittedly deals with rather abstract notions but these are, after all, linguistic universals, not extra-terrestrials.

If, by the way, the skill to distinguish lexical verbs from qualities on the one hand and from entities on the other indeed gradually develops autonomously, or through test wiseness, this is a very good reason to seriously doubt whether recognizing verbs, or whatever part of speech, indeed entails some sort of academic linguistic knowledge.

Be this as it may, the main point is that the results of the experiment proper render it impossible to demonstrate that School Grammar may very well instruct students effectively in recognizing parts of speech, but that this knowledge of grammar does not transfer to writing. After all, these results seem to point out rather decidedly that a course on School Grammar does not at all teach grammar effectively. And indeed, it doesn't, whereas in all probability a true traditional school grammar course, with semantics at its core, does. The Sentence Construction course itself is best proof for this. The course is built on *Grammatica in Balans*, a course heavily drawing on the true traditional school grammar. And it is no mere accident that the Sentence Construction course, or, more correct here, *Grammatica in Balans*, drew on these fundamentally semantic notions. But does that amount to saying that any traditional grammar course would yield what the Sentence Construction course yielded, or yet more? No.

Recall that the Sentence Construction course diverged rather sharply from *Grammatica in Balans* or whatever true traditional school grammar course. It left aside all that comes along traditionally in the school grammar curriculum and focused on what distinguishes sentences from non-sentences, on what distinguishes bare sentences from stuffed ones (also cf. chapter 4.4). Thus the course provides students with what they need to know; with knowledge they are able to use and indeed do use while writing, even if there is no overt cue that they should do so. A course on traditional school grammar, whatever its superior quality, will never do so as it does not touch upon what writers need to know.

8.3 The rise and fall of sentence awareness

The hypothesis pertaining to sentence awareness is that Sentence Construction students know what distinguishes sentences from non-sentences better than School Grammar and Direct Writing students do. Thus, it continues, they will use more genuine sentences in their writing than do School Grammar and Direct Writing students. Rather puzzlingly, results brought out that all students, no matter what they were taught, are experts on what distinguishes sentences from non-sentences. Still, though, and even more puzzlingly, the Sentence Construction students turn out to be the only ones that put this knowledge to use.

How come results bring out that students of either the School Grammar or the Direct Writing course, neither of them supplying explicit instruction on what distinguishes the fully-fledged sentences of written language from the utterances that sail under sentence colours in speech, do not seem to be particularly bothered by this negligence? How come their performance equals that of the students who were so instructed (cf. chapter 7.3, tables 7.7 and 7.8)?

Section 2, referring to chapter 6, already mentioned that whatever its effects, the effects of the School Grammar treatment are certainly treatment effects, and not fake effects brought about by whatever cause. The same goes for both other treatments. Neither was especially burdened by low-achievers or propelled by nonpareils. Both Direct Writing and Sentence Construction students were instructed in all they were meant to, in the way they were meant to, mastered all subject material they were meant to and did not confront anything that interfered either positively or negatively with this instruction. As tests nor analyses are defective, the results of the tests definitely should be taken to indicate that treatment is not what matters in sentence awareness.

So, contrary to the hypothesis, results seem to suggest that knowledge on sentences does come naturally to language users. Results of both follow-up studies (cf. tables 7.11 and 7.12) all the more illustrate the apparent autonomy of sentence awareness. Both paint the picture of students happily working away on tests that get easier and easier for them as they get older and older. In any

event, a course such as the Sentence Construction programme is not what students seem to need when it comes to gather such knowledge. Yet the colossal effect of the Sentence Construction programme on sentence awareness of students in both guided and non-guided writing (cf. figure 7.2) seems to suggest it is. The results of the extended experiment provide a bit of extra evidence for this conclusion.

So what's all this, then? Is it at all possible that the explicit instruction on what distinguishes sentences from non-sentences was not as superfluous as test results suggest? Could it be that the instruction was not wasted after all and that it did what it had to do: provide students with at least part of what they need to know about writing, viz. what are sentences? Or is the effect nothing but the spin-off of Miss repeatedly telling students after a tough instruction on sentences, patterns and completeness that a sentence is complete only when it starts with a capital letter and ends with a full stop? Remember that the Sentence Construction course is the only one dealing with the ins and outs of sentences. As both other courses seem to think sentences need not be bothered with, Direct Writing does not touch at all upon anything having to do with sentences, and School Grammar deals with punctuation accidentally only.

Could it be, then, that the due emphasis on punctuation, taught in tandem with the sentence patterns in the Sentence Construction course, had such an impact that students simply started sprinkling capitals and full stops over their writing, every so often hitting the mark? This definitely has a derogatory touch about it, suggesting students may have gone about punctuating rather unintelligently and as if not at all at home in the world of syntax and semantics, jotting down punctuation marks randomly. This is not what novices are inclined to do. They are not real monkeys. They may have started to partition off 'semantic and syntactic wholes' more frequently, though, thus scoring high on sentence ratio without the slightest inkling of a notion of what they were actually doing.

It is highly unlikely that this is what happened. First of all, for a sentence to be welcomed as such, it need not be punctuated at all. Sentence ratio is sensitized to all that a student might have possibly meant to be a sentence, no matter how deviantly stamped as such. There are no data on the various ways students segmented text into sentences but the young writers enrolled in the experiment certainly were a resourceful lot.

However, for a sentence to be a genuine sentence it need not just be demarcated in one way or the other. It is far more crucial for it to be a correct, untainted string of words with all its slots properly filled. T-unit ratio brings out that, not surprisingly, constructing syntactic strings poses no problem. Sentence ratio brings out (again: not surprisingly) that grouping these strings into sentences did. This is highlighted all the more by the obvious absence of any relation between T-unit ratio and sentence ratio (in appendix 2: $r=.55$, but be careful and take this for what is it worth; the correlation is grotesquely overestimated due to both linear and non-linear dependencies. Also cf. chapter 6.4.).

What are the problems novices confront in recognizing sentences in writing and did the Sentence Construction course in any way help diminish these problems?

Table 8.2 summarizes the essential changes treatment effected in sentence awareness in writing. It reveals that, irrespective of genre and conditions, at pre test one thirds of student-wise demarcated sentences was no sentence. In more than half of these erroneous sentences a word or two was missing, whereas almost one fifth was no sentence because there simply was no end to it.

Table 8.2. The essential changes treatment effected in sentence awareness of students in writing in the experiment proper (percentages) (DW = Direct Writing; SG = School Grammar; SC = Sentence Construction)

	Pre test	Post test		
		DW	SG	SC
Guided Description				
erroneous sentences, total	21	26	26	13
- sentences without end	1	0	0	0
- sentences missing a word (or two)	16	23	24	11
- rest and combination	4	3	2	2
Description				
erroneous sentences, total	39	43	35	34
- sentences without end	4	7	6	4
- sentences missing a word (or two)	29	22	22	20
- rest and combination	6	14	7	10
Narrative		35	39	24
erroneous sentences, total	31	15	14	6
- sentences without end	11	5	8	8
- sentences missing a word (or two)	7	15	17	10
- rest and combination	13			
Exposition				
erroneous sentences, total	42	45	43	32
- sentences without end	8	13	9	12
- sentences missing a word (or two)	17	13	12	9
- rest and combination	17	19	22	11
Overall				
erroneous sentences, total	33	37	36	26
- sentences without end	6	9	7	6
- sentences missing a word (or two)	17	16	17	12
- rest and combination	10	12	12	8

At post test, these percentages remain virtually unchanged for both School Grammar and Direct Writing students. Both student groups tend to demarcate sentences erroneously more frequently: Direct Writers do so for 37 of 100 sentences, School Grammar students for 36 of 100. One fifth of these would have been fine had there been a full stop; in about 40% an element is lacking. For Sentence Construction students percentages differ rather dramatically. Three quarters of all they think to be sentences are sentences indeed. As the table discloses, this is not because, good as gold, they took to punctuate meticulously. Rather, they took to writing full clauses. Overall, there is a 5% decrease in the number of incomplete sentences, not precisely a decrease to exult in but it is there all the same.

The effect is especially noticeable in both descriptive assignments. Perhaps it is safe to speculate that, under the influence of the course, emphasizing noun phrases are not sentences though they might very well seem to be so as they 'tell the same' (cf. chapter 4.4), listing in noun phrases tends to give way to listing in full clauses⁴. Thus, keen to write writing and possessed with some of the knowledge they need, Sentence Construction students know how to go about it. And so they do.

The same goes for the narrative genre. This one comes with its own problems: most erroneous sentences at pre test are non stop sentences. One might do well to combine this with the relatively high run-on ratio found in this genre: far over one third of all T-units, whether erroneous or not, starts with *and*, *but* or *then* (table 7.4). Taken together, these results conjure up in the mind's eye a vivid picture of an eager young narrator, carried away by his story's content, writing his head off, elaborating mainly by conjunction of what here may loosely be called propositions, and understandably under such circumstances, not especially heeding segmentation of prose. And yet, at post test, it seems Sentence Construction students shift their attention rather more frequently to the standard snatches a piece of prose such as this comes in. Apparently the spur of the treatment is enough to make them actually think over their writing.

Results of the extended experiment point out that this spur is not enough to make them do so for ever after. The effects ebb. They pop up again in the writing of ex-Direct Writers, though. This sustains the conclusion that the explicit instruction on what distinguishes sentences from non-sentences was not superfluous, whatever the results of both tests on sentence awareness suggest.

One might well wonder why, by the way, results of both tests do not at all herald the successfulness of the course in Sentence Construction. There is obviously more to recognizing sentences in writing than what is in these tests. This seems to be a matter of course: the tests point out a discrete problem to students, the writing assignments do not. If students possess any knowledge at all to solve the discrete problem, the tests are apt to bring this out. Writing need not: there, knowledge needs to transfer without the cue the tests supply.

Come to that, the discrete problem the tests present to students is no real problem at all, or so results say. Students easily discriminate sentences from non-sentences. So much so that at the very end of the experiment test reliability tumbles down: the tests get far too easy.

But, if all students are possessed of the knowledge they need, why is it that Sentence Construction students are the only ones to use it? Junior writers, differentiating written from oral language, desperately seek for all scraps of knowledge that might turn their writing into real writing. So why is it only Sentence Construction students do so whereas the rest do not?

4. The ratio of noun phrases per sentence should have provided details, here. But this ratio metamorphosed into a twin sister to sentence ratio in exactly the same way ratio of noun phrases per T-unit transformed into the double of T-unit ratio. A shame, really.

Recall that table 8.2 brought out that there are plenty of erroneous sentences. But then, the table disclosed there are plenty of non-erroneous strings as well. If, overall, one third of sentences is flawed, two thirds are flawless. So, then, how big a problem is recognizing sentences really? And could it be that some are more difficult to recognize as sentences than others?

It seems likely that this is the case. Remember that flawless sentences have a mean length of about four to seven words (cf. table 7.10). They are decidedly short, that is. So are the sentences students are asked to recognize or construct in both tests. So one might well speculate that short sentences are no problem at all, neither in tests nor in writing.

Results on sentence ratio in Description corroborate this conclusion. Notice that scores on sentence ratio for Description seem to be diametrically opposed to its scores on T-unit ratio (cf. table 7.2 and 7.9). Sentence ratio scores on Description are highest of the three impromptu genres (irrespective of condition or test period: .51, vs .44 for Narrative and Exposition), T-unit ratio scores are lowest (irrespective of condition or test period .73 vs .86 for Narrative and .78 for Exposition). Writing fluency in Description seems to be especially troubled by flawed strings. The majority of these turned out to be noun phrases: 'flawed' in that they were not visibly attached to an 'S'. Sentence ratio brings out that what flawless strings there are (and there are lots of these) are for the greater part correctly demarcated as genuine sentences. How come?

As was mentioned in the first section of this chapter, students quite rightly took descriptions for what they were: no ordinary prose but lists. Two aspects that distinguish a list from all other prose seem to be especially relevant here. First of all, a list fundamentally is a series of elements clad in simple structures such as noun phrases or simple statements (cf. (8.1-3), (2.2-3) and (2.8)). Second, lists in general do not look like prose texts: all noun phrases or simple statements start on a new line. In this study, anything that starts on a new line is taken to be possibly meant as a sentence. Some of it, then, is not. Most of it (the simple statements) is, though. So, short, simple sentences certainly seem no stumbling block for students. Hence their terrific test scores.

What are stumbling blocks then? Table 8.2 brings out that sentences are flawed for two reasons, mainly: either they are far too long or there are words missing. Both descriptive genres prominently miss words. This is largely due to the noun phrase effect. Not all of it is, though: sentences in narrative could do with some extra words as well and this genre is very emphatically not one that evokes lists. Rather it evokes information on the when and whereabouts of events in the story. Such information typically comes as a modification to sentences. It seems safe to speculate that these modifications, no sentences themselves, are likely to be recognized as such. Though they are afterthoughts of a kind, they are syntactic and semantic wholes and may even be conceptualized apart from the sentence they are a part of, all the more signalling their proclivity to stand alone (also cf. (2.10) to (2.12)).

This, of course, is crystal glazing. Section 1 already mentioned there is no detailed information as to what elements are missing from strings. The same goes here. But here it seems to be more than a mere bit of a pity.

Notice, by the way, that if this proclaimed tendency towards autonomy of sentence modifiers indeed is a problem for novices, treatment affects it in the expository genre only. All of them do; but Sentence Construction does most. The latter course abundantly deals with the sentential slots for extra information, highlighting the amount of precision, detail or interest these constituents add to the sentence without changing its basic pattern. This pays, apparently. As the tests do not question these problem sentences, it is impossible for them to bring out differences between conditions.

The same goes for the other problem in sentence awareness in writing, i.e. the writer's urge to carry on without end, especially in narratives. Studentwise demarcation of sentences discloses that some junior writers blurt out a narrative in one and the same breath, that others spurt, draw breath and spurt again and that some, finally, go forward step by step. Spurting or blurting of course contradicts the writer's urge to differentiate writing from speech. But then, perhaps, students may be quite confident when writing a narrative as the genre is so evidently familiar to them (cf. chapter 2.2). If there is any truth in this at all, table 8.2 brings out that the Sentence Construction course, as suggested earlier on, works a small wonder here: it makes long and short distance runners alike stop and think. Remember this is not an effect of the emphasis on capitals and full-stops. Merely jotting down punctuation marks is not half what comes into it.

As the tests do not question these problem sentences either, it seems to be no wonder they do not bring out the differences between conditions that junior writing attest to.

In short, though results of the extended experiment bring out its effects do not last, as they should have, the Sentence Construction programme seems to know how to handle the essential problems of sentence awareness rather well. Obviously the problem of sentence awareness is slightly more intricate than the tests supposed it to be. The tests are quite invalid, actually. So what about the first part of the hypothesis on sentence awareness stating that Sentence Construction students know what distinguishes sentences from non-sentences better than School Grammar and Direct Writing students do? It might hold after all. For, obviously, Sentence Construction students are possessed of knowledge other students go without.

8.4 Conciseness and overall writing quality

The hypothesis on the effects of the Sentence Construction course on conciseness of students turned out to be faulty for three quarters of its claims. When it came to deliberately squash information in noun phrases, the course met its match in the course on Direct Writing. Both surpassed the School Grammar programme. Not in writing, though. No treatment affected conciseness in writing.

Might it be the expectations of the effect of the Sentence Construction course in concise writing was a wee bit overstrung? After all, framing noun phrases firmly resides in the language faculty and for a grammar course to affect this would seem a bit thick. But then, attuning them to put to use what they are supposed to know tacitly, or should it say: prod them to exploit a simple trick, does not seem to be such a high aim. For surely, students know the trick but perform it very infrequently. Table 7.16 brought out that the amount of complex noun phrases in junior writing is definitely very small.

The smallest amount of complex noun phrases is to be found in the narratives, most are to be found in descriptions, whereas in the latter genre the mean total number of noun phrases is least, in Narrative most (6.89 and 13.89 respectively). This, of course, is because narratives were almost thrice the length of the descriptions (cf. table 7.1). As noun phrases were such a beloved gang in Description, the amount of noun phrases in narratives is not thrice the amount of noun phrases in description. In both descriptions and narratives, then, readers generally will stumble on two complex noun phrases at the utmost. Two types are frequent and as far as complexity goes, none too intricate. One type starts with a determiner, followed by one adjective, followed by a noun, the other is a determiner/noun combination, followed by a prepositional phrase (8% and 5% of complex noun phrases respectively). Simple noun phrases, with an article and a noun, are by far the most popular, especially in the narratives. There, 50 to 60 percent of all noun phrases are of the determiner-noun type whereas in Description (guided and free) and Exposition alike this is 35 percent only. A quarter to a third of all noun phrases are of the complex type in both descriptions.

As a means to get students to ponder over what they were writing and how they were putting it while they were at it, the Sentence Construction course trained students in adding information to genuine, albeit short, sentences. The course, then, had students deliberately squash a relatively large amount of information in relatively complex noun phrases. Here, it is not enough if students go by their tacit grammatical knowledge only: explicit knowledge on writing, i.e.: knowing that in writing this is what can be done, comes into it as well.

In this respect the Sentence Construction course failed. It succeeded in teaching students how to go about constructing complex noun phrases when a test (*viz.* Informing Succinctly) had them do so. That is: the course helped them solve the problem when it confronted students in discrete terms. The course, then, seems to have been on its way well enough but not far enough.

Yet one might wonder whether the Sentence Construction treatment is the best there is in order to get students to write concisely. The effects of the Direct Writing course reveal a course of a completely different stamp yields the same results. In the writing course, what could have been the ingredient that made students successfully inform succinctly? The most likely candidate seems to be

5. For it was the course and the course only that brought about this results, as has been brought forward time and again.

the spur experimenters continuously gave students to give more details, to 'tell more about this' (cf. chapter 4.3). Such an undifferentiated incitement to students to elaborate upon undeveloped ideas might have caused their attention to shift to tightly packing noun phrases. There were no such spurs when they worked on their test essays: these were 'impromptu' after all. Obviously the test provided enough of a spur.

The writing assignments could have been accompanied by a spur, by the way. If there would have been a clue as to whom the prose juniors wrote was meant for, results might have been quite different (cf. chapter 2.6). It seems regrettable no such cue was given. Or is it? On the one hand one could argue that a cue was absent quite rightly. After all, the Sentence Construction course was meant to initiate junior writers into the art of creating typical writing, elaborate and concise at the same time. If it had done so successfully, its students would have created typical writing, wouldn't they? And the answer to this is of course that yes, they would, so no, the course is not successful here. But then, might one not argue that writing for an undifferentiated public is hardly likely to result in differentiated prose (which is of course exactly what Kroll (1985a and b) and Smith and Combs (1980), cited in chapter 2.6, contend). Could it be that Sentence Construction students would outperform their peers because they are possessed of expert knowledge, shifting their attention from unawares producing structures of a certain type to exploiting these structure-types, making the most of the possibilities these vouchsafe the writer? This is a matter of mere speculation that should be engaged in very, very coolly. After all, none of the results suggest that the Sentence Construction students started to exploit sentences as yet. They are aware of sentences in writing, but they do not use them to structure content at this low level. Sentences have not become their building blocks. Or should it say: not yet? Perhaps this is a second step students were prepared, of preparing, to make. For, it seems, the course gave them the impetus towards spontaneous sentence awareness in writing (the first step) and towards informing succinctly if prompted to do so.

It remains open to debate whether Direct Writers, equal in skill to their Sentence Construction colleagues at informing succinctly, would have ever started to use sentences this way. The course does not seem to prompt them to do so. 'Telling more' afterwards seems to be a way of tackling writing essentially different to an approach that focuses on the sentence primarily. 'Telling more' implies the sentence is there already. Information is added afterwards only, reducing the generative power of the sentence itself to nil. The Sentence Construction approach to sentences is of a very different flavour altogether. While constructing sentences along the lines of the Sentence Construction programme, sentences grow 'on line', content begging for structure, structure signalling where extra content could go to.

Tackling writing this way would seem to result in quality prose. No juniors, however, tackled writing this way. Hence, quality of their prose remained unaffected. Thus, contrary to hypothesis, the Sentence Construction students did

not write prose of a quality superior to that of Direct Writers' prose. Neither did the latter's prose outstrip the quality of the prose of the School Grammar students.

As the Sentence Construction course succeeded in boosting students's sentence awareness in writing only, this is no surprise. Sentences are of no small importance in writing, but their being there, no matter how unconventionally clad, is by no means enough to positively affect quality. Not with writers this age, that is⁶.

It came as a bit of a disappointment that the Direct Writing course has nothing to show for, though. Students worked like mad, most of them enjoyed the course enormously, and, most important of all, they were reported to be more and more deliberate in their doings, evaluating the writing of others critically, heatedly defending their own prose. It seems as though there is not one teeny-weeny trace of the course to be found in the test essays students wrote. Could it be that students have a brainstorm only if Miss starts one? Could it be distance, criticism and wit is brought about in class conferences solely and not in non-discussed, non-read, non-revised writing? What about the quality of the prose these students wrote in class? Nothing, as yet. This matter surely needs to be studied in more depth.

8.5 Aftermath

By demonstrating that explicit, declarative knowledge on genuine sentences, taught to nine- and ten-year-olds in the a-rhetorical context of grammar class spontaneously transfers to writing, this study refutes both the claim that grammar instruction is of no use when it comes to ameliorate writing of juniors as well as the belief that the sort of grammar instruction traditionally associated with grammar in primary school is a good means to ameliorate writing. Thus, this study seems to qualify both assertions: explicit instruction on grammatical peculiarities of written language can help eradicate specific troubles in writing. In other words: a specialized grammar course, on constructing sentences, provides novice writers with specialized knowledge. This knowledge gets stored efficiently, to be used whenever the writer realizes it has to be used and, eureka, novice writers so provided use their knowledge. So far so good.

Questions of all kinds of flavours follow suit, though. A very prominent one seems to be the one that asks: what about this striking phenomenon? Does it actually suggest that, while writing, novices not merely search their memory for conscious knowledge on written language but also know where to find it? Or does it suggest the knowledge itself signals its usefulness while the head it is stored in is busy writing? Or what? Notions such as 'transfer' and 'monitoring' are relevant, here, but, though they were touched upon in this report, they

6. Cf. chapter 6.3, where someone ate a hat for this!

were not at its heart. Such questions will be at the heart of a next study, one that tries to picture when, where and perhaps how, explicit, conscious knowledge comes into writing.

Next in line, and probably closely related to the questions above, is the question as to the actual power of a programme styled like the Sentence Construction programme. Is it at all possible to penetrate its workings? How did it do what it seems to have been doing? And why did it not do what it was supposed to do, too? And, now that the contours of some of its vital ingredients can be guessed at more educatedly, couldn't the programme be made to be more effective, still?

Here, then, is to the results of yet another study that lifts yet another corner of the veil that shrouds writing and its development. Some questions seem answered, but, as most answers give rise to yet new questions, lots of questions remain unanswered.

At the start of the eighties, Wesdorp loudly vented his discontent with writing research in general and with studies into the effect of grammar instruction upon writing in particular (Wesdorp 1981, 1982, 1983; also cf. chapter 1). In reviewing research it came to him that he was looking over a motley crew, an archipelago of variegated theories, studies and results, all defective and pathetic in one way or the other. But then, he admitted, doing nothing would be no good either. So, he commended modesty and explicitness to the attention of researchers embarking upon enterprises such as the one reported on here. And what have we got here? Another study erratically focusing on a pygmy treatment, carrying on blindly, albeit resourcefully? This final résumé perhaps helps to settle the matter.

REFERENCES

- Algemene Nederlandse Spraakkunst* (1984) Geerts, G., Haeseryn, W., Rooy, J. de, & Toorn, M.C. van den (eds.) Groningen: Wolters Noordhoff.
- Bartholomae, D. (1980) The study of error. *College Composition and Communication* 31, 253-269.
- Battus (1981) *Opperlandse taal- en letterkunde*. Amsterdam: Querido
- Bereiter, C. & Scardamalia, M. (1982) From conversation to composition: The role of instruction in a developmental process. In R. Glaser (ed.), *Advances in Instructional Psychology*, Vol. 2. Hillsdale, N.J.: Lawrence Erlbaum Associates
- Bereiter, C. & Scardamalia, M. (1987) *The psychology of written language*. Hillsdale, N.J.: Lawrence Erlbaum Associates
- Bergh, H. van den & Kuhlemeier, H. (1989) *De proefpeiling Nederlands: een onderzoek naar de haalbaarheid van peilingsonderzoek in het Voortgezet Onderwijs*. Arnhem: Cito.
- Berken, N. van, Drost, W., Kramer C. et al. (1978) *Jouw Taal Mijn Taal*. Baarn: Bekadidakt.
- Bimbaum, J.C. (1982) The reading and composing behavior of selected fourth- and seventh-grade students. *Research in the teaching of English* 16, 241-260.
- Blok, H. (1987) *Taal voor alledag. Feiten en meningen over het taalgebruik van LBO- en MAVO-leerlingen in alledaagse situaties*. 's-Gravenhage: SVO.
- Bock, K. & Loebell, H. (1990) Framing sentences. *Cognition* 35, 1-39.
- Bos, D.J., & Oostdam, R.J. (1985) *Vorstudie periodieke peiling van het onderwijsniveau. Deel III: Het onderwijsaanbod Nederlands in de zesde klas*. Amsterdam: SCO.
- Bowey, J.A. (1986a) Syntactic awareness in relation to reading skill and ongoing reading comprehension monitoring. *Journal of experimental child psychology* 41, 282-299.
- Bowey, J.A. (1986b) Syntactic awareness and verbal performance from preschool to fifth grade. *Journal of psycholinguistic research* 15, 285-308.
- Braams, R. (1989) *Is zinsopbouwonderwijs te moeilijk en te abstract voor negenjarige basisschoolleerlingen?* Doctoraalscriptie Instituut voor Neerlandistiek. Universiteit van Amsterdam.
- Braddock, R., Lloyd-Jones, R., & Schoer L. (1963) *Research in written composition*. Champaign, Illinois: NCTE.
- Britton, J. (1970) *Language and learning*. Harmondsworth: Penguin.
- Brooks, L.Y. (1975) *The effect of a study of generative rhetoric on the syntactic fluency of seventh graders*. Doctoral Dissertation, Auburn University. AAD76-10890.
- Brus, B.Th., & Bergen, J.B.A.M. van (1979) *Schriftelijke opdrachten. Schoolvorderingentest voor het begrijpend lezen (bestemd voor het vierde leerjaar van de basisschool)*. Nijmegen: Berkhout.
- Burgess, C., Burgess, T. & Cartland, L. e.a. (1973) *Understanding children writing*. Harmondsworth: Penguin.

- Chafe, W.L. (1986) Writing in the perspective of speaking. In C.R. Cooper and S. Greenbaum (eds.) *Studying writing: linguistic approaches*. Beverly Hills: Sage Publishers.
- Christensen, F. & Christensen, B. (1978) *Notes toward a new rhetoric*. New York: Harper & Row.
- Cordeiro, P. (1988) Children's punctuation: an analysis of errors in period placement. *Research in the teaching of English* 22, 62-74.
- Daiute, C. (1985) Do writers talk to themselves? In: S.W. Freedman (ed.) *The acquisition of written language: response and revision*. Norwood, N.J.: Ablex Publishing Corporation.
- Daiute, C. (1989) Play as thought: thinking strategies of young writers. *Harvard Educational Review* 59, 1-23.
- Danielewicz, J. & Chafe, W. (1985) How 'normal' speaking leads to erroneous punctuation. In: S.W. Freedman (ed.) *The acquisition of written language: response and revision*. Norwood, N.J.: Ablex Publishing Corporation.
- Dirksen, A. (1990) *Monitoring processes in language production*. Doctoral Dissertation, University of Utrecht, Holland.
- Dort-Slijper, M.K. van (1984) *Grammatica in het basisonderwijs*. Leiden: Martinus Nijhof.
- Dort-Slijper, M. van, Gein, J. van de, & Vos, M. de (1988) *Grammatica in Balans*. Gorinchem: De Ruiter.
- Dyson, A.H. (1983) The role of oral language in early writing processes. *Research in the teaching of English* 17, 1-30.
- Edelsky, C. (1983) Segmentation and punctuation: developmental data from young writers in a bilingual program. *Research in the teaching of English* 17, 135-156.
- Elley, W., Barham, I., Lamb, H. & Wyllie, M. (1979) *The role of grammar in a secondary school curriculum*. Urbana, Ill.: NCTE.
- Faigley, L.L. (1979) The influence of generative rhetoric on the syntactic maturity and writing effectiveness of college freshmen. *Research in the teaching of English* 13, 197-206.
- Felix, S.W. (1987) *Cognition and language growth*. Dordrecht: Foris.
- Flower, L., Hayes, J.R., Carey, L. et al. (1986) Detection, diagnosis and the strategies of revision. *College Composition and Communication* 37, 16-55.
- Ford, M. & Holmes, V.M. (1978) Planning units and syntax in sentence production. *Cognition* 6, 35-53.
- Galambos, S.J. & Goldin-Meadow, S. (1990) The effects of learning two languages on levels of metalinguistic awareness. *Cognition* 34, 1-56.
- Geest, A.N. van der, & Swüste, W. (1974) *Spelenderwijs Stellen*. Nijmegen: Malmberg.
- Gein, J. van de (1989) *Effecten van grammatica-onderwijs op aspecten van stelvaardigheid van basisschoolleerlingen. Theorie, onderwijsprogramma's en toetsen*. Amsterdam: SCO.
- Gein, J. van de (1990) *Effecten van grammatica-onderwijs op aspecten van stelvaardigheid van basisschoolleerlingen. Deelrapportage resultaten II: kwaliteit van de uitvoering van het onderwijsexperiment*. Amsterdam: SCO.

- Glopper, K. de (1988) *Schrijven beschreven. Inhoud, opbrengsten en achtergronden van het schrijfonderwijs in de eerste vier leerjaren van het voortgezet onderwijs*. 's-Gravenhage: SVO.
- Graves, R. (1975) An examination of the writing processes of seven year old children. *Research in the teaching of English* 9, 227-241.
- Graves, D.H. (1980) When children want to punctuate: basic skills belong in context. *Language Arts* 57, 567-573.
- Grobe, C. (1981) Syntactic maturity, mechanics and vocabulary as predictors of quality ratings. *Research in the teaching of English* 15, 75-85.
- Hakes, D.T. (1980) *The development of metalinguistic abilities in children*. Berlin: Springer Verlag.
- Hardaway, J.M. (1969) *Generative rhetoric: an analyses of its influence on the writing of college freshmen*. Doctoral Dissertation, Arizona State University AAD69-20768.
- Haswell, R.H. (1981) Within-group distribution of syntactic gain through practice in sentence-combining. *Research in the teaching of English* 15, 87-96.
- Haswell, R.H. (1988a) Error and change in college student writing. *Written communication* 5, 479-499.
- Haswell, R.H. (1988b) Dark Shadows: the fate of writers at the bottom. *College Composition and Communication* 39, 303-315.
- Herriman, M.L. (1986) Metalinguistic awareness and the growth of literacy. In: S. Castell et al. (eds) *Literacy, society and schooling. A reader*. Cambridge: Cambridge University Press
- Hillocks, jr. G. (1986) *Research on written composition. New directions for teaching*. Urbana, Illinois: ERIC/RCE.
- Hunt, K.W. (1965) *Grammatical structures written at three age levels*. Champaign, Illinois: NCTE.
- Hunt, K.W. (1970) Syntactic maturity in school children and adults. *Monographs of the society for Research in Child Development* 35 (1). Chicago, Illinois.
- Kaufner, D.S., Hayes, J.R. & Flower, L. (1986) Composing written sentences. *Research in the teaching of English* 20, 121-140.
- Kerek, A., Daiker, D.A. & Morenberg, M. (1980) Sentence combining and college composition. *Perceptual and motor skills* 51, 1059-1157.
- Kinghorn, N.D., Faigley, L. & Clemens, Th. (1981) *A syntactic approach to college writing: an analysis of theory and effect*. University of North Dakota. (EDnr: 208 407).
- Krashen, S. (1977) The monitor model for adult second language performance. In: M. Burt, H. Dulay & M. Finocchiaro (eds.) *Viewpoints on English as a second language*. New York: Regents.
- Krashen, S. (1978) Individual variation in the use of the monitor. In: W.C. Ritchie (ed.) *Second language acquisition research. Issues and implications*. New York: Academic Press
- Kroll, B.M. (1981) Developmental relationships between speaking and writing. In B.M. Kroll & R. Vann (eds.) *Exploring speaking-writing relationships: connections and contrasts*. Urbana, Illinois: NCTE.

- Kroll, B.W. (1985a) Rewriting a complex story for a young reader: the development of audience adapted writing skills. *Research in the teaching of English* 19, 120-139.
- Kroll, B.W. (1985b) Social-cognitive ability and writing performance: how are they related? *Written Communication* 2, 293-305.
- Loban, W.D. (1963) *The language of elementary school children*. Champaign, Illinois: NCTE.
- Loban, W. (1976) *Language development: Kindergarten through grade twelve*. Urbana, Illinois: NCTE.
- Martin, W. (1980) *The effects of a program of models-imitation on the writing of seventh grade students*. Doctoral Dissertation, University of Austin, Texas. AAD81 00938.
- Matsushashi, A. (1981) Pausing and planning. The tempo of written discourse production. *Research in the teaching of English* 15, 113-134.
- McCutchen, D. (1988) 'Functional Automaticity' in children's writing. A problem of metacognitive control. *Written Communication* 5, 306-324.
- Mellon, J.C. (1969) *Transformational sentence-combining; A method for enhancing the development of syntactic fluency in English composition*. Champaign, Illinois: NCTE.
- Naber, L., Nijhof, S. & Schute, A. (1986) *Analyse van taalmethoden*. Den Bosch: Katholiek Pedagogisch Centrum.
- Neilsen, L. & Piché, G. (1981) The influence of headed nominal complexity and lexical choice on teacher's evaluation of writing. *Research in the teaching of English* 15, 65-74.
- Nold, E.W. (1982) Revising: intentions and conventions. In: R.A. Sudol (ed.) *Revising: New essays for teachers of writing*. Urbana, Illinois: ERIC/NCTE.
- Nold, E.W. & Freedman, S.W. (1977) An analysis of reader's responses to essays. *Research in the teaching of English* 11, 164-174.
- O'Donnell, R.C., Griffin, W.J. & Norris R.C. (1967) *Syntax of kindergarten and elementary school children: a transformational analysis*. Champaign, Illinois: NCTE.
- O'Hare, F. (1973) *Sentence-combining; improving student writing without formal grammar instruction*. Urbana, Illinois: NCTE.
- Perera, K. (1984) *Children's writing and reading. Analysing classroom language*. London: Basil Blackwell.
- Perera, K. (1986) Grammatical differentiation between speech and writing in children aged 8 to 12. In A. Wilkinson (ed.) *The writing of writing*. Philadelphia: Open University Press.
- Plompen, P. (1990) *De didactiek van het zelfstandig naamwoord in Jouw Taal Mijn Taal, een synthetische taalmethode voor de basisschool*. Doctoraalscriptie Instituut voor Neerlandistiek, Universiteit van Amsterdam.
- Potter, R.R. (1976) Sentence structure and prose quality: an exploratory study. *Research in the teaching of English* 1, 17-28.
- Purcell-Gates, V. (1988) Lexical and syntactic knowledge of written narrative held by well-read-to Kindergartners and second graders. *Research in the teaching of English* 22, 128-160.

- Quirk, R. & Greenbaum, S. (1973) *A university grammar of English*. London: Longman.
- Rosen, C., & Rosen, H. (1973) *The language of primary school children*. Harmondsworth: Penguin.
- Scardamalia, M., Bereiter, C. & Goelman, H. (1982) The role of production factors in writing ability. In: N. Nystrand (ed.) *What writers know: The language, process and structure of written discourse*. San Diego: Academic Press Inc.
- Schuurs, U. (1990) *Leren schrijven voor lezers. Het effect van drie vormen van probleemgericht schrijfonderwijs op de zinsbouwvaardigheid*. Enschede: Faculteit Wijsbegeerte en Maatschappijwetenschappen.
- Shaughnessy, M.P. (1977) *Errors and expectations. A guide for the teacher of basic writing*. New York: Open University Press
- Sinclair, H., Jarvella, R.J. & Levelt, W.J.M. (1978) *The child's conception of language*. Berlin: Springer Verlag.
- Smith, G.L. (1982) Revision and improvement: making the connection. In: R.A. Sudol (ed.) *Revising: New essays for teachers of writing*. Urbana, Illinois: ERIC/NCTE.
- Smith, W.L. & Combs, W.E. (1980) Effects of overt and covert cues on written syntax. *Research in the teaching of English* 14, 19-38.
- Spiegel, D.L. & Fitzgerald, J. (1990) Textual cohesion and coherence in children's writing revisited. *Research in the teaching of English* 24, 48-66.
- Stevens, R., Madden, N., Slavin, R. & Farnish, A. (1987) Cooperative integrated reading and composition: two field experiments. *Reading Research Quarterly*, 22, 433-454.
- Steward, M.F. & Grobe, C.H. (1979) Syntactic maturity, mechanics of writing and teacher's quality ratings. *Research in the teaching of English* 13, 207-215.
- Strong, W. (1986) *Creative approaches to sentence combining*. Urbana, Illinois: ERIC/NCTE.
- Sulzby, E. (1986) Young children's concepts for oral and written text. In K. Durkin (ed.) *Language development in the school years*. London: Croom & Helm.
- Tordoir, A. & Damhuis, R. (1982) *Grammatica en vreemde-talenonderwijs in de brugklas. Een onderzoek naar de aansluiting van het moedertaalgrammatika-onderwijs op het vreemde-talenonderwijs*. Amsterdam: SCO.
- Tordoir, A. & Wesdorp, H. (1979) *Het grammatica-onderwijs in Nederland. Een researchoverzicht betreffende de effecten van grammatica-onderwijs en een verslag van een onderzoek naar de praktijk van dit onderwijs in Nederland*. 's-Gravenhage: SVO.
- Tunmer, W.E. & Bowey, J.A. (1984) Metalinguistic awareness and reading acquisition. In: W.E. Tunmer, C. Pratt & M.L. Herriman (eds.) *Metalinguistic awareness in children. Theory, research and implications*. Berlin: Springer Verlag.
- Tunmer, W.E. & Herriman, M.L. (1984) The development of metalinguistic awareness: a conceptual overview. In W.E. Tunmer, C. Pratt & M.L. Herriman (eds.) *Metalinguistic awareness in children. Theory, research and implications*. Berlin: Springer Verlag.

- Watts, R.W. (1980) *The effects of grammar-free sentence-combining practice on the syntactic fluency of the written and oral responses of seventh grade students*. Doctoral dissertation University of South Carolina. AAD81-02792.
- Weaver, C. (1982) Welcoming errors as signs of growth. *Language Arts* 59, 438-444.
- Wesdorp, H. (1981) *Evaluatietechnieken voor het moedertaalonderwijs. Een inventarisatie van beoordelingsmethoden voor de stelvaardigheid, het begrijpend lezen, de spreek-, luister- en discussievaardigheid*. 's-Gravenhage: SVO.
- Wesdorp, H. (1982) Over de problemen bij toetsingsonderzoek op het gebied van het moedertaalonderwijs. *Tijdschrift voor Taalbeheersing* 4, 22-41.
- Wesdorp H. (1983) *Schrijven in het voortgezet onderwijs*. 's-Gravenhage: SVO.
- Wesdorp, H. (1984) The evaluation of the traditional grammar curriculum in The Netherlands. *Studies in educational evaluation*, 10, 283-297.
- Williams, J.D. (1987) Covert linguistic behavior during writing tasks: psychophysiological differences between above-average and below-average writers. *Written Communication* 9, 310-328.
- Zwarts, M. (ed.) (1990) *Balans van het taalonderwijs aan het einde van de basisschool. Uitkomsten van de eerste taalpeiling einde basisonderwijs*. Arnhem: Cito.

APPENDICES

- Appendix 1. Rules of thumb for preparing essays for analysis: making chunks
- Appendix 2. Correlations (p.m.c.) between four tests and nine measures, summed over four writing tests
- Appendix 3. Correlations (p.m.c.) between measures for Guided Description, Description, Narrative and Exposition

Appendix 1. Rules of thumb for preparing essays for analysis: making chunks

1. All studentwise demarcated sentences are the beginning of a chunk. A chunk is either a flawless or a flawed T-unit or is a sentence fragment. If it is a sentence fragment, it might either make a (flawed or non-flawed) T-unit with the preceding and/or following fragment or stand alone.
2. Chunks beginning with *en* (Eng.: *and*), *maar* (Eng.: *but*), *want* (Eng.: *for or as*) and *toen* (Eng.: *then*)¹ are chunked right in from of their openers, but not if the string they open is reduced.
3. If an error in a reduced main clause is caused by the absence of the grammatical subject, the clause-without-subject is a new chunk.
4. If a noun phrase is not part of a coordination it is a chunk unless the author announces a list.
5. Direct speech is one chunk up to the point where a new chunk starts.
6. Exclamations, stereotyped expressions, answers, dialogues etc. are separate chunks.
7. Discontinuous strings are represented the way the author wrote them. Numbers signal there is a chunk within a chunk.
8. Sounds are separate chunks unless they are part of a string.
9. Titles at the head of essays and literally repeated text openers plus topic are not rated. Chunk number: 00; text in square brackets.

1. Toen the adverb, not the conjunction (which is, in English: *when*).

Appendix 2. Correlations (p.m.c.) between four tests and nine measures, summed over four writing tests (N=148)

*	rec.verbs	T-unit	ml	spl	ror	con.sen	rec.sen	sen.rat	ml	spl	inf.suc	com.op	qual.sc
recognizing verbs	1.00												
T-unit ratio	.25	1.00											
mean T-unit length	.25	.69	1.00										
noun phrases/T-ratio	.27	.94	.69	1.00									
run-on ratio	-.19	.02	.23	.22	1.00								
construct sentences	.28	.26	.22	.26	-.16	1.00							
recogn. sentences	.27	.36	.24	.35	-.20	.49	1.00						
sentence ratio	.17	.55	.33	.52	-.03	.15	.32	1.00					
mean sentence length	.20	.47	.62	.47	.12	.22	.28	.72	1.00				
noun phrases/sentence	.17	.53	.33	.53	.01	.14	.31	.99	.73	1.00			
informing successfully	.28	.31	.33	.30	-.14	.31	.40	.33	.36	.33	1.00		
compl. noun phr. ratio	-.01	.14	.29	.14	.13	.07	-.02	.09	.24	.11	.13	1.00	
overall quality score	.24	.39	.49	.39	.10	.36	.32	.28	.41	.28	.38	.24	1.00

Academic linguistic knowledge

Recognizing Verbs

Writing fluency

- main measure
- T-unit ratio
 - exploratory measures
 - Mean T-unit length
 - Noun phrases per T-unit ratio
 - Run-on ratio

Sentence awareness

- Constructing Sentences
Recognizing Sentences

- main measure
- Sentence ratio
 - exploratory measures
 - Mean sentence length
 - Noun phrases per sentence

Coherence

- Informing Successfully
Complex noun phrase ratio

Overall quality

- overall quality score

Appendix 3. Correlations (p.m.c.) between measures for Guided Description, Description, Narrative and Exposition (N=148)

*	tr-gd	mtl-gd	sr-gd	msl-gd	cr-gd	tr-des	mtl-des	ror-des	sr-des	msl-des	cr-des	glob-des	tr-nar
tr-gd	1.00												
mtl-gd	.81	1.00											
sr-gd	.99	.80	1.00										
msl-gd	.26	.72	.24	1.00									
cr-gd	.19	.48	.21	.61	1.00								
tr-des	.25	.17	.24	.02	.01	1.00							
mtl-des	.24	.20	.23	.08	-.01	.66	1.00						
ror-des	.10	.12	.09	.10	.04	.04	.21	1.00					
sr-des	.13	.15	.15	.08	.08	.56	.28	.08	1.00				
msl-des	.18	.21	.19	.13	.83	.48	.52	.06	.73	1.00			
cr-des	-.07	-.05	-.06	.03	-.08	-.05	.18	.11	.01	.23	1.00		
glob-des	.20	.27	.27	.24	.21	.27	.33	.12	.26	.27	.02	1.00	
tr-nar	.09	.07	.09	-.05	.01	.06	-.02	.01	.14	.13	-.11	.08	1.00
mtl-nar	.34	.31	.32	.11	.05	.24	.31	.09	.07	.18	-.01	.21	.39
ror-nar	.06	.07	.05	.11	.09	-.08	-.02	.21	-.13	-.10	-.05	-.07	.02
sr-nar	.10	.14	.09	.06	.08	.22	.09	-.00	.50	.25	-.10	.21	.27
msl-nar	.19	.26	.18	.17	.11	.08	.11	.08	.27	.22	-.06	.26	.20
cr-nar	.14	.06	.13	-.06	.03	.13	.16	.03	.02	.02	.03	.01	.03
glob-nar	.31	.43	.31	.37	.23	.13	.14	.02	.21	.32	.04	.42	.15
tr-exp	.14	.13	.14	.08	.13	-.10	.04	-.11	.05	.13	.07	.02	.13
mtl-exp	.10	.15	.10	.15	.16	.03	.12	.04	.17	.32	.09	.13	.17
ror-exp	.09	.14	.06	.13	.17	-.03	.10	.15	.01	-.05	.05	.21	.06
sr-exp	.11	.05	.13	-.01	.00	.12	.00	.01	.45	.31	-.01	.11	.18
msl-exp	.13	.15	.15	.11	.05	.12	.14	.07	.31	.39	.12	.15	.18
cr-exp	.22	.17	.22	.04	.11	.03	.00	.13	.01	.07	-.02	.11	.06
glob-exp	.24	.34	.23	.23	.27	.15	.17	.09	.11	.16	.01	.41	.09

cont.

-continued-

*	mtl-nar	ror-nar	sr-nar	msl-nar	cr-nar	glob-nar	tr-exp	mtl-exp	ror-exp	sr-exp	msl-exp	cr-exp	glob-exp
mtl-nar	1.00												
ror-nar	.31	1.00											
sr-nar	.05	-.01	1.00										
msl-nar	.42	.20	.67	1.00									
cr-nar	.17	-.02	-.03	.08	1.00								
glob-nar	.41	-.09	.16	.33	-.03	1.00							
tr-exp	.07	.03	.01	.04	-.02	.15	1.00						
mtl-exp	.25	.01	-.07	.07	-.03	.34	.57	1.00					
ror-exp	.09	.16	.06	.16	.05	-.01	.04	.07	1.00				
sr-exp	.06	.02	.40	.22	.07	.35	.33	-.06	.100	1.00			
msl-exp	.25	.01	.16	.18	.02	.37	.63	-.02	.75	.100	1.00		
cr-exp	.09	-.01	.05	.01	-.01	.20	-.04	.11	-.03	.11	.12	1.00	
glob-exp	.25	-.09	.08	.17	.06	.48	.24	.19	.18	.00	.08	-.02	1.00

- tr-gd = T-unit ratio Guided Description
 mtl-gd = Mean T-unit Length Guided Description
 sr-gd = Sentence ratio Guided Description
 msl-gd = Mean Sentence Length Guided Description
 cr-gd = Complex noun phrase ratio Guided Description
 tr-des = T-unit ratio Description
 mtl-des = Mean T-unit ratio Description
 ror-des = Run-on ratio Description
 sr-des = Sentence ratio Description
 msl-des = Mean Sentence Length Description
 cr-des = Complex noun phrase ratio Description
 glob-des = Overall essay quality Description
 tr-nar = T-unit ratio Narrative

- mtl-nar = Mean T-unit ratio Narrative
 ror-nar = Run-on ratio Narrative
 sr-nar = Sentence ratio Narrative
 msl-nar = Mean Sentence Length Narrative
 cr-nar = Complex noun phrase ratio Narrative
 glob-nar = Overall essay quality Narrative
 tr-exp = T-unit ratio Exposition
 mtl-exp = Mean T-unit ratio Exposition
 ror-exp = Run-on ratio Exposition
 sr-exp = Sentence ratio Exposition
 msl-exp = Mean Sentence Length Exposition
 cr-exp = Complex noun phrase ratio Exposition
 glob-exp = Overall essay quality Exposition

Author's index

ANS 22

Bartholomae 21

Bereiter & Scardamalia 12, 14, 15,
17, 18, 20

Bergh, Van den & Kuhlemeier 21

Bergh, Van den & Eiting 92

Birnbaum, 12, 14, 18

Blok 21

Bock & Loebell 13, 20

Bos 4

Bowey 10

Braams 90 (fnt)

Braddock et al. 22, 23

Britton 11

Brooks 28

Burgess et al. 11

Chafe 20

Christensen 27, 28, 31, 33

Cordeiro 19

Daiute 15, 17, 18

Danielewicz & Chafe 20

Dirksen 3, 4

Dort-Slijper 47

Dyson 10

Edelsky 19

Eiting 96 (fnt)

Elley et al. 23

Faigley 29

Felix 10

Flower et al. 18

Ford & Holmes 20

Galambos & Goldin-Meadow 10

Gein 63, 81 (fnt)

Glopper 21

Grammatica in Balans 40, 52-60,
82, 126-127

Graves 11, 19

Grobe 76

Hakes 10

Hardaway 28

Haswell 21, 26

Herriman 8, 10

Hillocks 2, 22, 26, 29, 76

Hunt 13, 16, 67

Jouw Taal Mijn Taal 40, 47-52, 125

Kaufner et al. 18

Kerek et al. 25, 76

Kinghorn 19

Krashen 19,

Kroll 16, 17, 27, 134

Loban 13, 16

Martin 23, 26

Matsuhashi 20

McCutchen 15

Mellon 24-26, 76

Naber et al. 47

Neilsen & Piché 77

Nold 15

Nold & Freedman 76

O'Donnell et al. 13, 16

O'Hare 25, 26, 76

Perera 14, 16, 17

Plompen 52 (fnt)

Potter 76

Purcell-Gates 10

Quirk & Greenbaum 22

Rosen & Rosen 11

- Scardamalia et al. 12-15, 17, 18, 20
Schuurs 3, 4
Shaughnessy 21
Sinclair et al. 10
Smith 14
Smith & Combs 26, 27, 134
Spelenderwijs Stellen 40, 44-47
Spiegel & Fitzgerald 76
Stevens et al. 22, 23
Steward & Grobe 76
Strong 26, 76
Sulzby 10

Tordoir & Damhuis 1
Tordoir & Wesdorp 1
Tunmer & Bowey 10
Tunmer & Herriman 10

Watts 26
Weaver 17
Wesdorp 1, 2, 136
Williams 18

Zwarts 21

Subject's index (variables, tests and measures in the study at hand come in small capitals)

ACADEMIC LINGUISTIC KNOWLEDGE 34, 60,
71-72, 104-106, 124-127
a-rhetorical setting 4, 24-25, 33, 56,
120, 135
attention/control/monitor 3, 8-9, 10-
12, 14-19, 23, 26, 29-31, 34-36,
121, 123, 130, 134-136
complexity (syntactic)/syntactic
integration 8, 10, 26-29, 30-31,
61, 69, 76, 77, 78, 79, 93, 97,
133-135
CONCISENESS 35-36, 46, 60, 75-76,
113-115, 132-135
control
- of nuisance variables 5, 36,
41, 66, 96, 119
- conscious attention/monitor 3,
8-9, 10-12, 14-19, 23, 26, 29-
31, 34-36, 121, 123, 130, 134-
136
conversation/conversation-like/con-
versational 7-9, 12, 14-15, 20,
29-31, 121
correlations 1 (motto), 94-96, 128
- overall quality x syntactic
quality 92-93
demarcation/punctuation/segmenta-
tion of sentences 19-22, 34, 48,
52, 59, 128, 129-131
differentiation 15-19, 29-30, 34,
130, 132, 134
DIRECT WRITING 34-36, 44-47, 60-61,
87-90, 124, 126, 128, 129, 132-
135

discrete point tests
- pilot study 64
- RECOGNIZING VERBS 71-72, 104-
106, 124-126
- RECOGNIZING SENTENCES 72-75,
106, 110, 127-132
- CONSTRUCTING SENTENCES 72-75,
106, 110, 127-132
- INFORMING SUCCINCTLY 75-76,
113, 132-135
effects
- ceiling effects 64, 92
- experimenter effects 37, 91
- genre effects 67, 96, 123, 130
- memory effects 72
- school effects 36
- topic effects 66
- of follow up treatment 84
- of generative rhetoric/sentence
construction 27, 29
- of miscellaneous grammar in-
struction 22-24
- of sentence combining 24-27
experimenters 59-60, 86-87
experimenter effects 37, 91
finite form (Du: persoonsvorm) 48-
49, 124-125
frames
- sentence frames 14, 31, 51,
52, 55, 56, 59
- syntactic frames 13, 22, 30,
52, 55, 123
- conceptual frame 13, 20, 94
generative rhetoric/sentence con-
struction 27-29, 52, 76
grammar instruction (traditional
school ~) 1-4, 6, 34-35, 40, 52,
124-127, 135-136

- grammatical knowledge
- tacit 3-4, 9, 10, 18, 22, 30, 121, 133
 - explicit/declarative 3, 23, 29, 34-35, 126, 133, 135
 - expert/specialized 4, 18, 30, 34, 61, 121, 127-132, 134, 135
- integration (syntactic)/syntactic complexity 8, 10, 26-29, 30-31, 61, 69, 76, 77, 78, 79, 93, 97, 133-135
- kernel of the sentence (Du: zinskern) 48, 124
- mean T-unit length 67, 69-70, 71, 95, 103
- monitor/attention/control 3, 8-9, 10-12, 14-19, 23, 26, 29-31, 34-36, 121, 123, 130, 134-136
- overall quality of written language 22, 24-25, 29, 76 (also see: written language)
- OVERALL WRITING QUALITY 35-36, 60-61, 67, 76-77, 115-116, 132-135
- parsing sentences and recognizing parts of speech 34, 49, 52, 124
- persoonsvorm (Eng: finite form) 48-49, 124-125
- pilot study
- tests and measures 64-67
 - treatments 43-44
- protocol 43
- PROTOCOL ADHERENCE 38-39, 87
- punctuation/demarcation/segmentation of sentences 19-22, 34, 48, 52, 59, 128, 129-131
- punctuation marks 9, 19-22, 67, 70, 74, 94-95, 128, 129, 131-132
- REGULAR CURRICULUM 40, 90-91
- SCHOOL GRAMMAR 34-36, 47-52, 60-61, 88-90, 124-127, 128, 129, 132, 135
- segmentation/punctuation/demarcation of sentences 19-22, 34, 48, 52, 59, 128, 129-131
- SENTENCE AWARENESS 34-36, 64, 72-75, 106-113, 127-132
- sentence combining 24-27, 76
- SENTENCE CONSTRUCTION 34-36, 52-59, 88-90, 124-127, 127-132, 132-135
- sentence construction/generative rhetoric 27-29, 52, 76
- sentence patterns 52, 54, 55, 56, 59
- SENTENCE RATIO 74-75, 106-110, 110-111
- STUDENTS' PROGRESS 39-40, 89-90
- syntactic maturation/syntactic fluency/syntactic growth 24-29
- T-unit 67-70, 74, 121
- T-UNIT RATIO 68-70, 71, 74, 94-95, 102, 121-124
- TIME ON TASK 39, 88-89transfer 3-4, 23, 29-30, 34, 126, 130, 135
- writing assignments 64-67
- WRITING FLUENCY 67-71, 101-104, 121-124
- written language 3-4, 7-9, 10, 15-19, 21, 22, 29-31, 34-35, 60, 72, 127-135 (also see: overall quality of written language)
- zinskern (Eng: kernel of the sentence) 48, 124

SAMENVATTING (Summary in Dutch)

Proefleider a op school 5:
"Waar begint een complete zin mee?"
Leerling: "Met het eerste woord."

Proefleider c op school 6:
"Wie weet nog wat een zinsdeel is?"
Leerling: "Limonade."

Inleiding

Het onderzoek naar de 'effecten van grammatica-onderwijs op aspecten van stelsvaardigheid van basisschoolleerlingen', zoals deze studie voluit heet, kan beschouwd worden als een vervolg op de studie die Wesdorp uitvoerde naar de doelen en resultaten van het grammatica-onderwijs. Een eerdere poging strandde door gebrek aan theorie; Wesdorp trof uitsluitend overtuigingen aan en die, meende hij, konden niet de theoretische basis leggen voor fatsoenlijk effect-onderzoek. In deze studie gaat het erom aannemelijk te maken dat het traditionele schoolse grammatica-onderwijs geen heilzame invloed heeft op de stelsvaardigheid van negen- à tienjarig schrijvers. Dit komt niet, zo luidt de veronderstelling, doordat het grammatica-onderwijs is, maar doordat het de verkeerde soort grammatica-onderwijs is. Het schools-traditionele grammatica-onderwijs leert kinderen niet wat ze moeten weten om beter te schrijven. Dan kunnen de leerlingen tijdens het schrijven dus ook onmogelijk gebruik maken van de schools-traditionele kennis.

Om hun stelsvaardigheid te verbeteren is het overigens wel noodzakelijk dat leerlingen vertrouwd gemaakt worden met enkele grammaticale kenmerken van schrijftaal. Hun onbewuste grammaticale kennis voorziet daar immers niet in. Dat betekent echter beslist niet dat kinderen moeten leren wat wel en niet welgevoerde syntactische taalstructuren zijn. Zelfs de jongste schrijvertjes, hoe onervaren ook, weten op hun duimpje welke structuren wel en welke niet welgevoerd zijn. Doordat hun pen hun gedachten onmogelijk kan bijhouden, zullen ze tijdens het schrijven weleens één of twee noodzakelijke woordjes vergeten. Doordat ze hun tekstjes niet herlezen, zien ze misschien nooit dat ze een zin hebben opgeschreven die een beetje raar in elkaar zit. Hun aandacht zal af en toe verslappen en dat kan bijvoorbeeld maken dat twee zinsstructuren in één zin samensmelten.

Wellicht experimenteren ze met schrijftaal en produceren ze daardoor onbeholpen zinnen met deftige woorden. De vaardigheid om correct te formuleren is echter rotsvast in het menselijk taalvermogen verankerd en wat dat betreft is er dan ook niets mis.

Mis is het met het zinsbenul. De zin, althans: de geschreven volzin, is namelijk geen natuurlijke eenheid in taal. Voor schrijvers is de volzin echter een onmisbare kapstok: het is in de tekst het laagste niveau waar vorm en inhoud elkaar ontmoeten. Leerlingen moeten derhalve vertrouwd gemaakt worden met zinnen. Ze moeten leren welke zinsdelen tenminste in een zin aanwezig moeten zijn en ze moeten leren hoe ze aan zo'n basiszin informatie kunnen toevoegen. Hier komt het begrippenapparaat om de hoek kijken dat in de regel geassocieerd wordt met het grammatica-onderwijs.

In zeker opzicht heeft dit onderzoek raakvlakken met de studies van Schuurs en Dirksen. Zij onderscheiden ongelukjes (A-fouten) en interactionele fouten (I-fouten) in opstellen van twaalfjarigen. De ongelukjes, het woord zegt het al, zijn eigenlijk helemaal geen fouten; men hoeft zich er ook geen zorgen over te maken. Als een schrijver zijn opstel nog eens overleest, zal 'ie de ongelukjes er zonder veel problemen uithalen. Daar zorgt zijn onbewuste grammaticale kennis voor. Die onbewuste kennis rekent echter niet af met fouten die erop terug te voeren zijn dat de schrijver te weinig rekening houdt met de lezer voor wie hij schrijft. Schuurs en Dirksen meenden dat grammatica-onderwijs kinderen zou leren deze fouten te vermijden. De resultaten van hun onderzoek lijken hun geen gelijk te geven.

Dit onderzoek draait ook om fouten die met de onbewuste grammaticale kennis niet opgelost kunnen worden: het draait om de zin. Grammatica-onderwijs zal leerlingen leren wat in schrijftaal een zin is. De resultaten van dit onderzoek maken aannemelijk dat dat inderdaad zo is.

Wat schrijvers moeten weten

Centraal in het onderzoek naar de effecten van grammatica-onderwijs op teksten van jonge schrijvers is de overtuiging dat grammatica-onderwijs leerlingen attent kan maken op de voornaamste kenmerken van de vreemde taal waarmee ze vertrouwd moeten raken: standaard schrijftaal. Die taal verschilt in een aantal opzichten van spreektaal. Eén van de meest opvallende verschillen is de aanwezigheid van zinnen in schrijftaal.

Dit onderzoek wil ook laten zien dat leren wat in schrijftaal een zin is niet noodzakelijkerwijs in de context van het stelonderwijs hoeft te gebeuren. De hypothese luidt dat effecten van een zinsopbouw cursus zich in teksten van leerlingen zullen manifesteren zonder dat leerlingen in die cursus ooit hebben geschreven. Van de specialistische zinskennis zal spontaan transfer plaatsvinden naar het stellen.

Dit kan gebeuren doordat leerlingen meer en meer de taal die zij spreken en de taal die zij schrijven van elkaar willen laten verschillen. Schrijven ze aanvankelijk nog zoals ze spreken, als ze wat ouder worden doen leerlingen welbewuste

pogingen echte schrijftaal te produceren. Zo laten hun tekstjes bijvoorbeeld zien dat ze rekening proberen te houden met het doel van de tekst en met het publiek waar de tekst voor geschreven is. Ze schrijven ook zinnen waarvan de bouw complexer is dan van doorsnee spreektaalzinnen.

Schrijven van schrijftaal kost onervaren schrijvers echter moeite. Ze weten immers niet precies wat zoal de kenmerken van schrijftaal zijn. Ze hebben er slechts een vermoeden van. Dat vermoeden leidt hen. Soms leidt het hen de goede kant op, soms de verkeerde. Zekerheid is er nooit. Het ontbreekt onervaren schrijvers immers aan kennis. Daardoor weten zij werkelijk niet wat ze moeten doen.

Een illustratie van dit ontbreken van kennis is het zinsbenul waarvan leerlingen in hun tekstjes blijf geven. Zodra kinderen hebben leren schrijven en ook met hoofdletters en punten kennis gemaakt hebben, beginnen zij te interpungeren. Aanvankelijk doen zij maar wat: ze schrijven bijvoorbeeld elk woord met een hoofdletter en zetten er een punt achter of beginnen elke nieuwe regel met een hoofdletter en zetten achter elke regel een punt, of daar nu een nieuwe zin begint of niet. Uiteindelijk beginnen ze eenheden af te bakenen die naar structuur en betekenis afgerond zijn. Het lijkt erop dat zij zich daarbij oriënteren op de zogenaamde 'intonation unit': een samenhangende uiting die uit ongeveer vijf tot zeven woorden bestaat, waarvan het uitspreken zo'n twee seconden duurt en die kenmerkend is voor spreektaal. Dit zijn echter niet de eenheden die voor schrijftaal kenmerkend zijn: schrijftaalzinnen zijn in de regel veel langer dan uitingen in spreektaal. Leerlingen lijken dus zinsdelen voor zinnen aan te zien. Dat is niet een probleem waarmee alleen negen- en tienjarigen kampen. Afgaande op de interpunctie hebben ook vijftienjarige en eerstejaars studenten een gebrekkig zinsbesef.

De experimentele cursussen

Om leerlingen zinsbesef bij te brengen is een grammaticacursus ontworpen. Het idee daarvoor is ontleend aan de 'generative rhetoric' van Christensen. In dat programma staat het bouwen van zinnen centraal. Leerlingen leren een onderscheid te maken tussen de basiszin, die elementaire informatie bevat, en de extra informatie die daaraan toegevoegd kan worden. Ze worden attent gemaakt op de mogelijkheden die de basiszin hun geeft om extra informatie toe te voegen. Zo leren ze nadenken over vorm en inhoud van de basiszin en over inhoud en plaats van extra informatie. Ze overwegen dus, terwijl ze schrijven, welbewust wat ze schrijven en hoe ze dat schrijven. Dit heeft een positief effect op de kwaliteit van teksten.

De grammaticacursus, of liever: zinsopbouw cursus ('Sentence Construction') is gebaseerd op materiaal uit *Grammatica in Balans*. In *Zinnen Bouwen*, zoals het leerlingboek genoemd is, worden leerlingen vertrouwd gemaakt met zinspatronen. Die patronen, gefundeerd op de woordsoorten en afhankelijk van de subcategorisatie van het werkwoord, hangen nauw samen met de betekenis van de zin. Vorm

en betekenis worden zo in hun onderlinge relatie behandeld. Centraal staat, net als in het Christensen-programma, de basiszin en de mogelijkheden die deze biedt voor verlevendiging en precisering van de zinsinhoud.

Effecten van zo'n zinsopbouw cursus worden in dit onderzoek gecontrasteerd met effecten van een cursus traditionele schoolgrammatica en met de effecten van een cursus waarin leerlingen uitsluitend schrijven. Van beide wordt in het basis-onderwijs immers verwacht dat ze de stelvaardigheid van leerlingen zullen verbeteren. De cursus traditionele schoolgrammatica ('School Grammar'), waarvan het leerlingmateriaal *Zinnen en delen van zinnen* genoemd is, is ontleend aan *Jouw Taal Mijn Taal*, een veel gebruikte, moderne synthetische taalmethode die relatief grondig aandacht aan grammatica besteedt. In *Zinnen en delen van zinnen* leren leerlingen de kern van de zin afbakenen. Ze leren ook zinnen ontleden en woorden benoemen.

De stelcursus ('Direct Writing'), tenslotte, is samengesteld met behulp van *Spelenderwijs Stellen*, een kistje vol losse opdrachtkaarten. In deze cursus worden leerlingen vertrouwd gemaakt met verschillende tekstgenres. Ze leren ook hoe ze het schrijven van een tekst het beste aan kunnen pakken: ze leren 'brainstormen', ze worden gewoon eerst een klad te schrijven, dat klad te herlezen en herschrijven en pas dan een net in hun werkboek te schrijven. Dat werkboek is *Werkboek Stellen* genoemd.

Alle cursussen zijn voorzien van draaiboeken ('protocols') waarin les voor les zeer gedetailleerd voorgeschreven is wat er wanneer en hoe gebeuren moet.

Alle drie de cursussen en de draaiboeken zijn uitgetprobeerd en daarna uitgebreid herzien.

Hypothesen

Leerlingen weten al wat in hun moedertaal wel en niet syntactisch welgevormde structuren zijn. Geen van de drie cursussen zal, of moet, of kan hun dat leren. De eerste hypothese in dit onderzoek luidt dan ook dat de formuleervaardigheid ('writing fluency') van leerlingen in orde is en onveranderd blijft.

Leerlingen moeten daarentegen wel leren wat in schrijftaal zinnen zijn. Dat leren ze noch van een cursus traditionele schoolgrammatica noch van een stelcursus. Die twee cursussen zullen dus ook geen positief effect hebben op het zinsbesef, hoewel de cursus traditionele schoolgrammatica wel zal maken dat leerlingen dezelfde schools-grammaticale kennis tot hun beschikking hebben als de zinsopbouwleerlingen. De tweede hypothese van het onderzoek luidt derhalve dat zinsopbouw- en school-grammaticaleerlingen evenveel schools grammaticale kennis ('academic linguistic knowledge') zullen hebben. Beide leerlinggroepen hebben, vanzelfsprekend, meer van zulke kennis dan de stellerleerlingen. De derde hypothese stelt echter dat alleen de zinsopbouw cursus maar een gunstig effect op het zinsbesef ('sentence awareness') van leerlingen zal hebben.

Leerlingen die over zinsbesef beschikken, weten ook welke hoekjes en gaatjes er in de zin zitten. Die plekjes zullen ze uitbuiten. Dat resulteert in de beknoptheid ('conciseness') die voor schrijftaal zo kenmerkend is. De zinsopbouwleerlingen onderscheiden zich ook van beide andere groepen leerlingen door deze beknoptheid, aldus de vierde hypothese.

Volgens de vijfde en laatste hypothese zal de globale kwaliteit ('overall writing quality') van de teksten van de zinsopbouwleerlingen die van die van de andere leerlinggroepen overtreffen. Alleen de zinsopbouwleerlingen overwegen immers bewust inhoud en vorm van hun zinnen en dus van hun tekstjes.

Testen en maten

Voor elk van de vijf afhankelijke variabelen zijn meetinstrumenten ontworpen. Sommige daarvan zijn discrete taken, andere zijn stelopdrachten. Beide zijn op grote schaal uitgeprobeerd. Dat resulteerde uiteindelijk in vier testjes en vier stelopdrachten.

Er is één test om de schools grammaticale kennis van leerlingen te meten (Werkwoorden Herkennen; 'Recognizing Verbs'), er zijn er twee voor het zinsbenul (Zinnen Bouwen en Zinnen Herkennen; 'Constructing Sentences' en 'Recognizing Sentences') en er is er eentje voor beknoptheid (Beknopt Informatie Geven; 'Informing Succinctly'). De psychometrische kwaliteit van de tests is voldoende gebleken.

Er is één gestuurde steltaak ('Guided Description') en er zijn drie open stelopdrachten. In de gestuurde taak moeten leerlingen telkens in één zin een beschrijvinkje geven. In de open stellingen maken leerlingen ook een beschrijving ('Description') van bijvoorbeeld een ijsje naar keuze, of een kinderkopie, zij schrijven een fantasieverhaal ('Narrative') over bijvoorbeeld de fluitketel die in het schoolorkest wou, of over de regenlaarzen die uit wandelen gingen en ze schrijven een uiteenzetting ('Exposition') over bijvoorbeeld hun beste vriendje of vriendinnetje of over hun favoriete boek.

Formuleervaardigheid van leerlingen wordt gemeten met de T-unit ratio: de verhouding van het aantal woorden dat in syntactisch welgevormde structuren staat en het totaal aantal woorden in het opstel. Deze maat is ontleend aan Hunt's T-unit, maar er is één belangrijke wijziging aangebracht. Traditioneel is een T-unit per definitie welgevormd: onwelgevormde T-units worden gerepareerd tenzij ze niet te repareren zijn; dan worden ze niet in een analyse betrokken. In deze studie, echter, bestaan ook onwelgevormde T-units. Er werden géén T-units gerepareerd en alle T-units werden in de analyses betrokken, wel- of niet welgevormd. T-unit ratio is psychometrisch onbetrouwbaar gebleken.

Het zinsbenul waarvan leerlingen in hun teksten blijken te geven wordt gemeten met de zinsratio: de verhouding van het aantal woorden dat in door de leerling gemarkeerde volzinnen staat en het totaal aantal woorden in het opstel. Zinnen hoeven niet per se regulier als zinnen gemarkeerd te zijn. Nummertjes voor of schuine streepjes tussen zinnen, elke zin op een nieuwe regel; als leerlingen maar

op één of andere manier laten zien dat ze weten wat een zin is. De zinsratio is een betrouwbare maat.

Beknoptheid in teksten is gemeten met de complexe nominale constituenten ratio ('complex NP-ratio'): de verhouding van het aantal complexe nominale constituenten en het totaal aantal nominale constituenten. Er zijn diverse typen en niveaus van complexiteit onderscheiden. De maat bleek onbetrouwbaar.

Al deze syntactische kenmerken van teksten zijn beoordeeld door negen beoordelaars, alle negen afgestudeerde of bijna afgestudeerde taalkundigen of taalbeheersers. Hun betrouwbaarheid was naar tevredenheid.

Globale kwaliteit is gemeten door een jury van telkens vier leden, alle vier ervaren basisschoolleerkrachten, een globaal kwaliteitsoordeel te laten geven over de beschrijvingen, de fantasieverhalen en de uiteenzettingen. Zowel de betrouwbaarheid van de maat als van de jury waren in orde.

Opzet van de studie

Het ontwerp van de studie naar de effecten van grammatica-onderwijs op teksten van jonge schrijvers kan losjes omschreven worden als een 'randomized pre to post test control group design'. Van elke zesde groep van de zes aan het experiment deelnemende scholen zijn de leerlingen a-select aan één der drie experimentele condities toegewezen. 'School' en 'conditie' zijn dus niet met elkaar gecontamineerd.

De lessen werden niet door de groepsleerkrachten maar door proefleiders ('experimenters') gegeven. De proefleiders, drie in totaal, gaven niet in maar één conditie les, maar in alle drie. 'Conditie' en 'proefleider' waren dus ook niet met elkaar gecontamineerd.

Het eigenlijke experiment liep gedurende dertien schoolweken. In die dertien weken kregen de leerlingen elke week twee maal één vol uur les. Ze werden drie maal getest: één maal voorafgaand aan het experiment, één maal halverwege en één maal aan het eind ervan.

Om ervoor te zorgen dat één van de drie experimentele cursussen aan alle leerlingen van één klas gegeven was, startte na het eigenlijke experiment een vervollexperiment. Dat duurde precies even lang als het eigenlijke experiment. Hoewel het in eerste instantie bedoeld was als dienstverlening aan de deelnemende scholen, was het een mooie gelegenheid om het effect te bestuderen van een bepaalde volgorde van cursussen. In het vervollexperiment zijn de leerlingen derhalve nog twee maal getest: halverwege en aan het eind. In het totaal hebben de leerlingen dus vijf maal tests en stelopdrachten moeten maken.

Het ontwerp van het experiment (ook wel bekend onder de naam 'randomized block design') beveiligd het onderzoek (en de onderzoeker) weliswaar voor enkele van de grote gevaren in experimenteel onderzoek, het garandeert toch nog niet dat de effecten die gevonden worden inderdaad veroorzaakt zijn door experimen-

tele interventie. Verschillen per conditie in proefleidergedrag, de tijd die leerlingen met de stof bezig zijn, de vorderingen van de leerlingen en het taalonderwijs dat zij van de eigen groepsleerkracht krijgen, kunnen ook heel goed verantwoordelijk zijn voor effecten. Of voor het uitblijven van effecten. Zodoende zijn gedurende de uitvoering van het experiment de draaiboekvastheid ('protocol adherence') van de proefleiders, time on task en vorderingen van de leerlingen ('students' progress') en het reguliere taalonderwijs ('regular curriculum') gecontroleerd. Er bleek nergens sprake van systematische verschillen tussen condities. Hierdoor is vast komen staan dat, welke effecten er ook zijn, het inderdaad experimentele effecten zijn.

Resultaten

Met de formuleervaardigheid van leerlingen is, zoals ook gesteld in de eerste hypothese, weinig mis. Naar het lijkt heeft inderdaad geen der drie cursussen er effect op gehad. Doordat T-unit ratio psychometrisch onbetrouwbaar is (naar alle waarschijnlijkheid door gebrek aan systematische variantie), kunnen effecten echter niet gemeten worden. Er lijkt sprake van een gering genre-effect: leerlingen gebruiken in een beschrijving minder welgevormde structuren dan in een fantasieverhaal. Dit is toe te schrijven aan de definiëring van de T-unit in deze studie. In een beschrijving zijn leerlingen geneigd op te sommen wat zij zien. Zulke beschrijvingen bestaan dus uit series naamwoordgroepen. Dat zijn weliswaar correcte syntactische structuren, maar geen volledige en dus onwelgevormde T-units.

De zinsopbouw cursus is de enige die effect heeft op de schools-grammaticale kennis van de leerlingen. Tegen de verwachting in kunnen de leerlingen die met de traditionele school-grammatica cursus gewerkt hebben niet beter woordsoorten herkennen dan de stelleren. Het lijkt erop dat dat komt doordat de cursus traditionele school-grammatica de semantische aspecten van de woordsoorten verwaarloost. Dit maakt dat leerlingen geen inzicht krijgen in het universele onderscheid tussen zelfstandigheden (de zelfstandige naamwoorden), activiteiten of werkingen (de werkwoorden) en hoedanigheden (de bijvoeglijke naamwoorden). Ze leren weliswaar de persoonsvorm herkennen, maar ze weten niet waar dat een persoonsvorm van is.

Het zinsbesef waarvan leerlingen in de twee discrete zinsbenultaken blijken geven doet vermoeden dat leerlingen heel best weten wat zinnen zijn. Beide testjes worden steeds beter gemaakt en er is geen enkel effect te bespeuren van welke cursus dan ook. In de tekstjes zijn het echter alleen de zinsopbouwleerlingen die hun zinsbenul gebruiken: zij overtreffen wat dat betreft zowel de stelleren als de traditionele grammaticale leerlingen. Dat betekent dus dat er van de grammaticale kennis over zinnen in schrijftaal inderdaad spontaan transfer plaatsvindt naar stelvaardigheid. Zonder dat de leerlingen daartoe aangemoedigd worden, doen zij

een beroep op hun zinskennis. Dat van dit superieure zinsbenul in de discrete taken niets blijkt, moet waarschijnlijk toegeschreven worden aan invaliditeit van die taakjes. Daarin blijken bij nader inzien nu juist niet die zinnen te staan die voor schrijvertjes zonder zinsopbouw kennis een probleem vormen als zij zelf een tekst maken.

Resultaten in het vervollexperiment laten zien dat het effect van de zinsopbouw cursus wegebt zodra leerlingen niet meer met de cursus werken: zinsopbouwleerlingen die daarna met de stelcursus werken zijn aan het einde van het vervollexperiment weer terug bij af. Bij leerlingen die eerst teksten hebben leren schrijven en daarna met de zinsopbouw cursus les krijgen, openbaart zich het effect wel weer.

Tegen de verwachting in laat de beknoptheidstest zien dat de stelcursus net zo'n gunstig effect heeft als de zinsopbouw cursus. 't Effect manifesteert zich ook in het vervollexperiment. De traditionele grammaticacursus heeft geen effect. Doordat beknoptheid in teksten niet betrouwbaar gemeten is, kunnen de effecten daar niet geanalyseerd worden. Dat de stelcursus zo'n gunstig effect heeft, hier, komt wellicht doordat de proefleiders, geheel volgens draaiboek, de schrijf-leerlingen steeds aanmoedigen méér informatie te geven, méér te vertellen in hun teksten. Het is mogelijk dat zo'n ongerichte proefleideraanmoediging leerlingen geattendeerd heeft op de mogelijkheden om meer informatie te geven en toch beknopt te zijn.

Er is geen sprake van dat toegenomen zinsbesef en beknoptheid van de zinsopbouwleerlingen ook maakt dat zij zo bewust vorm en inhoud van wat zij schrijven overdenken dat dit een gunstig effect heeft op de tekstkwaliteit. De globale kwaliteit van teksten, door wie ook geschreven, blijft onveranderd.

Tot slot

Het is evident dat zinsopbouwonderwijs maakt dat leerlingen zinsbenul verwerven en gebruiken. Kennelijk geeft de cursus ook de eerste aanzet tot grotere beknoptheid van leerlingen. Het volgende stapje, naar spontane beknoptheid, maken de zinsopbouwleerlingen echter niet. Of is het beter om te zeggen: nog niet? Voorlopig lijkt het erop dat daardoor ook de kwaliteit van hun teksten onveranderd blijft.

De stelcursus zet leerlingen ook aan tot grotere beknoptheid. Zij ontberen echter alle zinsbesef en detailleren en verlevendigen hun 'zinnen' pas achteraf. Hoewel hier de nodige speculatie om de hoek komt kijken, lijkt deze aanpak vrij fundamenteel te verschillen van de aanpak die de zinsopbouwleerlingen in alle waarschijnlijkheid ontwikkelen. Zinsopbouwleerlingen doen immers een voortdurend en spontaan beroep op hun zinsbenul en dat maakt dat zij uiteindelijk, zich bewust van vorm en inhoud, een staketsel bouwen, er inhoud bijbouwen en de hele boel zonodig weer verbouwen.

De resultaten van dit onderzoek laten zien dat het mogelijk is om stelvaardigheid van leerlingen positief te beïnvloeden met behulp van alleen maar grammatica-onderwijs. Zulk grammatica-onderwijs mag best volledig buiten de context van het schrijfcurriculum gegeven worden, maar het moet wel onderwijs zijn dat leerlingen voorziet van kennis die ze nodig hebben als ze schrijven. Van zulke kennis vindt spontaan transfer plaats naar stelvaardigheid.

Curriculum vitae

Jannemieke van de Gein was born 6 November, 1954 in Vlaardingen. She attended the Scholengemeenschap Spieringshoek at Schiedam (gymnasium alpha). She started studying English Language and Literature at Leiden University in 1976 but turned to general linguistics at Utrecht University after receiving her B.A. in 1979. She received her master's degree in general linguistics in 1982. As an immediate follow-up she was involved in working out alternative ways to instruct primary school students in grammar. As a result of this she was appointed assistant researcher to the linguistics section of the Dutch department at the University of Amsterdam in 1984. There, she helped developing a new grammar course for primary school.

From 1987 to spring 1991 she has been with the Research Group on Language Education at the Centre for Educational Research of the University of Amsterdam, working on the study into the effects of grammar instruction upon junior writing. In 1990 she took up lecturing in stylistics and applied linguistics at the Dutch department of Leiden University.